Large deviation principle of Freidlin-Wentzell type for pinned diffusion processes *

Yuzuru INAHAMA †

Abstract

Since T. Lyons invented rough path theory, one of its most successful applications is a new proof of Freidlin-Wentzell's large deviation principle for diffusion processes. In this paper we extend this method to the case of pinned diffusion processes under a mild ellipticity assumption. Besides rough path theory, our main tool is quasi-sure analysis, which is one of the deepest parts of Malliavin calculus.

1 Introduction

For the canonical realization of d-dimensional Brownian motion $(w_t)_{0\le t\le 1}$ and the vector fields $V_i : \mathbf{R}^n \to \mathbf{R}^n$ $(1 \le i \le d)$ with sufficient regularity, let us consider the following Stratonovich-type stochastic differential equation (SDE):

$$dy_t = \sum_{i=1}^{d} V_i(y_t) \circ dw_t^i \qquad \text{with} \quad y_0 = a \in \mathbf{R}^n.$$

For simplicity of explanation, no drift term is added, but modification is easy. The correspondence $w \mapsto y$ is called the Itô map and denoted by $y = \Phi(w)$. It is well-known that the Itô map is not continuous as a map from the Wiener space. Moreover, it is not continuous with respect to any Banach norm on the Wiener space which preserves the structure of the Wiener space.

Now, introduce a small positive parameter $\varepsilon \in (0, 1]$ and consider

$$dy_t^\varepsilon = \sum_{i=1}^{d} V_i(y_t^\varepsilon) \circ \varepsilon dw_t^i \qquad \text{with} \quad y_0^\varepsilon = a \in \mathbf{R}^n.$$

* **Mathematics Subject Classification:** 60F10, 60H07, 60H99, 60J60. **Keywords:** large deviation principle, pinned diffusion process, stochastic differential equation, rough path theory, quasi-sure analysis

† Graduate School of Mathematics, Nagoya University Furocho, Chikusa-ku, Nagoya 464-8602, JAPAN. E-mail: `inahama@math.nagoya-u.ac.jp`

Formally, $y^\varepsilon = \Phi(\varepsilon w)$. The process $(y_t^\varepsilon)_{0\le t\le 1}$ takes its values in $\mathbf{R}^n$ and its law is a diffusion measure associated with the starting point a and the generator $\mathcal{L}^\varepsilon = (\varepsilon^2/2)\sum_{i=1}^d V_i^2$.

A classic result of Freidlin and Wentzell states the laws of $(y_t^\varepsilon)_{0\le t\le 1}$ satisfies a large deviation principle as $\varepsilon \searrow 0$. The proof was not so easy. (See Friedman [4] or Dembo-Zeitouni [2] for instance). If Φ were continuous, we could use contraction principle and the proof would be immediate from Schilder's large deviation principle for the laws of $(\varepsilon w)_{0\le t\le 1}$. However, it cannot be made continuous in the framework of the usual stochastic analysis.

Ten years ago, Ledoux, Qian, and Zhang [14] gave a new proof by means of rough path theory, which was invented by T. Lyons [15]. See also [8, 16] for example. Roughly speaking, a rough path is a couple of a path itself and its iterated integrals. Lyons established a theory of line integrals along rough paths and ordinary differential equation (ODE) driven by rough paths. The Itô map in the rough path sense is deterministic and is sometimes called the Lyons-Itô map. The most important result in the rough path theory could be Lyons's continuity theorem (also known as the universal limit theorem), which states that the Lyons-Itô map is continuous in the rough path setting.Brownian motion (w_t) admits a natural lift to a random rough path W, which is called Brownian rough path. If we put W or εW into the Lyons-Itô map, then we obtain the solution of Storatonovich SDE (y_t) or (y_t^ε), respectively. They proved in [14] that the laws of εW satisfy a large deviation principle of Schilder type with respect to the topology of the rough path space. Large deviation principle of Freidlin-Wentzell type for the laws of (y_t^ε) is immediate from this, since contraction principle can be used in this framework. Since then many works on large deviation principle on rough path space have been published [3, 6, 7, 13, 19].

Then, a natural question arises; can one obtain a similar result for pinned diffusion processes with this method, too? More precisely, does the family of measures $\{\mathbb{Q}_{a,a'}^\varepsilon\}_{\varepsilon>0}$ satisfy a large deviation principle as $\varepsilon \searrow 0$? Here, $\mathbb{Q}_{a,a'}^\varepsilon$ is the pinned diffusion measure associated with $\mathcal{L}^\varepsilon$, which starts at a at time $t=0$ and ends at a' at time $t=1$. Heuristically, $\mathbb{Q}_{a,a'}^\varepsilon$ is the law of y_1^ε under the conditional probability measure $\mathbb{P}(\,\cdot\,|y_1^\varepsilon = a')$, where $\mathbb{P}$ stands for the Wiener measure.

The aim of this paper is to answer this question affirmatively under certain mild ellipticity assumption for the coefficient vector fields. Besides rough path theory, our main tool is quasi-sure analysis, which is a sub-field of Malliavin calculus. It deals with objects such as Watanabe distributions (i.e., generalized Wiener functionals) and capacities associated with Sobolev spaces. Recall that motivation for developing this theory was to analyse (the pullbacks of) pinned diffusion measures on the Wiener space.

Takanobu and Watanabe presented this kind of large deviation principle under a hypoellipticity assumption for coefficient vector fields (Theorem 2.1, [23]). This result seems very general and nice, but they gave no proof. Their tool are Malliavin calculus, and in particular, quasi-sure analysis. Recall that rough path theory did not exist, then. Presumably, they computed Besov norm of the solution of SDE, but details are unknown.

Since we use rough path theory, we will compute, not the output, but the input of the (Lyons-)Itô map. Here, the input means (w_t) itself and its iterated Storatonovich stochastic integrals. Hence, we believe that our proof via rough paths is probably simpler. Extending our method to the hypoellipticity case is an interesting and important future task.

Another preceding result is by Hsu [10] for a special case. He proved the case for (scaled) Brownian bridge on a complete Reimannian manifold M (i.e., the case $\mathcal{L}^{\varepsilon} = (\varepsilon^2/2)\triangle_M$, where $\triangle_M$ stands for Laplace-Beltrami operator on M). His proof is based on estimates and asymptotics for the heat kernel of $\triangle_M/2$ and no SDE appears in his paper. In this sense, this nice result of Hsu is not so "preceding" ours and it may not be very suitable to call it "Freidlin-Wentzell type".

The organization of this paper is as follows. In section 2, we give a precise setting, introduce assumptions, and state our main result (Theorem 2.1). In section 3, we introduce Besov-type norms on the geometric rough path space and prove their basic properties in relation to Hölder norms. In rough path theory, Hölder norms and variation norms are most important, while Besov norms merely play auxiliary roles. However, in connection with Malliavin calculus, Besov norms on the rough path space become very important. In section 4, we give a brief survey of Malliavin calculus and quasi-sure analysis for later use. We give basic facts on capacities and Watanabe distributions. Sugita's theorem is the most important among them, which states that a positive Watanabe distribution is actually a finite Borel measure on the abstract Wiener space.

In section 5, following Higuchi [9] and Aida [1], we recall that Brownian motion w admits a natural lift quasi-surely and that this version of Brownian rough path W is ∞-quasi-continuous. In section 6, using this quasi-continuity and Sugita's theorem, we obtain a probability measure on the geometric rough path space such that its pushforward by the Lyons-Itô map induces the pinned diffusion measure in question.

Sections 7–9 are the main ingredients of this paper. In Section 7, we consider the family of finite Borel measures on the geometric rough path space, which are the lift of the measures corresponding to positive Watanabe distributions $\delta_{a'}(y_1^{\varepsilon}) = \delta_{a'}(y^{\varepsilon}(1,a))$. In Theorem 7.1, which is a key result in this paper, we state a large deviation principle for them. Notice that, in the proof, ellipticity only at the starting point a is assumed. Since the Lyons-Itô map is continuous, the pushforward measures also satisfy a large deviation principle. This is a special case of Takanobu and Watanabe's result (Theorem 2.1, [23]). Our main theorem (Theorem 2.1) is almost immediate from this, when the pinned diffusion measures are well-defined and normalization is possible.

Sections 8 and 9 are devoted for the proof of Theorem 7.1. These two sections are the core part of our efforts. Several key facts used in these sections are as follows; (i) large deviation estimate for capacities on geometric rough path space, (ii) integration by parts formula in the sense of Malliavin calculus for Watanabe distributions, (iii) uniform non-degeneracy of Malliavin covariance matrix for solutions of the shifted SDE.

2 Setting and main result

In this section we give a precise setting and state our main result. Let $(w_t)_{0\le t\le 1}$ be the canonical realization of d-dimensional Brownian motion. We consider the following $\mathbf{R}^n$-valued Stratonovich-type SDE;

$$dy_t^\varepsilon = \sum_{i=1}^d V_i(y_t^\varepsilon)\circ \varepsilon dw_t^i + V_0(\varepsilon, y_t^\varepsilon)dt \qquad \text{with} \quad y_0^\varepsilon = a \in \mathbf{R}^n. \tag{2.1}$$

Here, $\varepsilon \in [0,1]$ is a small parameter and $V_i \in C_b^\infty(\mathbf{R}^n, \mathbf{R}^n)$ for $1\le i\le d$ and $V_0 \in C_b^\infty([0,1]\times\mathbf{R}^n, \mathbf{R}^n)$. (A function is said to be of class C_b^∞ if it is a bounded, smooth function with bounded derivatives of all order.) For each ε, (y_t^ε) is a diffusion process with its generator

$$\mathcal{L}^\varepsilon = \frac{\varepsilon^2}{2}\sum_{i=1}^d V_i^2 + V_0(\varepsilon,\,\cdot\,).$$

We assume everywhere ellipticity:

(H1): For all $a\in\mathbf{R}^n$, the set of vectors $\{V_1(a),\dots,V_d(a)\}$ linearly spans $\mathbf{R}^n$.

Under this assumption, the heat kernel $p_t^\varepsilon(a,a')$ exists and positive for all $a,a'\in\mathbf{R}^n$, $t>0$ and $\varepsilon>0$. Hence, the pinned diffusion measure $\mathbb{Q}_{a,a'}^\varepsilon$ associated with $\mathcal{L}^\varepsilon$ exists for any $\varepsilon>0$, starting point a and terminal point a'. It is a unique probability measure on $C_{a,a'}([0,1],\mathbf{R}^n) := \{x\in C([0,1],\mathbf{R}^n) \mid x_0 = a, x_1 = a'\}$ characterized by the following equation; for any $l\in\mathbf{N}_+$, any partition $\{0=t_0<t_1<\cdots<t_l<t_{l+1}=1\}$ of $[0,1]$, and any bounded, continuous function $f:\mathbf{R}^{n\times l}\to\mathbf{R}$,

$$\int f(x_{t_1},\dots,x_{t_l})\mathbb{Q}_{a,a'}^\varepsilon(dx) = p_1^\varepsilon(a,a')^{-1}\int_{(\mathbf{R}^n)^l} f(a_1,\dots,a_l)\prod_{j=1}^{l+1} p_{t_j-t_{j-1}}^\varepsilon(a_{j-1},a_j)\prod_{j=1}^{l} da_j.$$

Here, $a_0 = a$ and $a_{l+1}=a'$ by convention. In fact, $\mathbb{Q}_{a,a'}^\varepsilon$ sits on

$$C_{a,a'}^{\alpha-H}([0,1],\mathbf{R}^n) = \{x\in C_{a,a'}([0,1],\mathbf{R}^n) \mid x \text{ is } \alpha\text{-Hölder continous }\}$$

for any $\alpha\in(1/3,1/2)$. Heuristically, $\mathbb{Q}_{a,a'}^\varepsilon$ is the law of y_1^ε under the conditional probability measure $\mathbb{P}(\,\cdot\,|y_1^\varepsilon = a')$, where $\mathbb{P}$ stands for the Wiener measure. (This argument can be made rigorous with quasi-sure analysis, however.)

Let $\mathcal{H}$ be Cameron-Martin space for (w_t). For $h\in\mathcal{H}$, we denote by $\phi^0=\phi^0(h)$ be a unique solution of the following ODE;

$$d\phi_t^0 = \sum_{i=1}^d V_i(\phi_t^0)dh_t^i + V_0(0,\phi_t^0)dt \qquad \text{with} \quad \phi_0^0 = a. \tag{2.2}$$

We set $\mathcal{K}^{a,a'} = \{h \in \mathcal{H} \mid \phi^0(h)_1 = a'\}$, which is not empty under **(H1)**.

Define a rate function $J : C^{\alpha-H}_{a,a'}([0,1], \mathbf{R}^n) \to [0, \infty]$ by

$$J(y) = \inf\{\frac{\|h\|^2_{\mathcal{H}}}{2} \mid h \in \mathcal{K}^{a,a'} \text{ with } y = \phi^0(h)\} - \min\{\frac{\|h\|^2_{\mathcal{H}}}{2} \mid h \in \mathcal{K}^{a,a'}\}$$

if $y = \phi^0(h)$ for some $h \in \mathcal{K}^{a,a'}$ and define $J(y) = \infty$ if no such $h \in \mathcal{K}^{a,a'}$ exists. In fact, J will turn out to be good.

Now we state our main result in this paper. The rest of the paper will be devoted to proving this theorem.

Theorem 2.1 *Let $1/3 < \alpha < 1/2$ and assume* **(H1)**. *The family $\{\mathbb{Q}^{\varepsilon}_{a,a'}\}_{\varepsilon>0}$ of probability measures on $C^{\alpha-H}_{a,a'}([0,1], \mathbf{R}^d)$ satisfies a large deviation principle as $\varepsilon \searrow 0$ with a good rate function J, that is, for any Borel subset $A \subset C^{\alpha-H}_{a,a'}([0,1], \mathbf{R}^n)$,*

$$-\inf_{y \in A^\circ} J(y) \le \liminf_{\varepsilon \searrow 0} \varepsilon^2 \log \mathbb{Q}^{\varepsilon}_{a,a'}(A) \le \limsup_{\varepsilon \searrow 0} \varepsilon^2 \log \mathbb{Q}^{\varepsilon}_{a,a'}(A) \le -\inf_{y \in \bar{A}} J(y).$$

3 Geometric rough path space with Besov norm

In this section we introduce the geometric rough path space with Besov norm and recall its relation to the one with Hölder norm. In rough path theory, we usually use Hölder and variation norms. In connection to Malliavin calculus, however, Besov norm play an essential role. Chapter 7 and Appendix A.2 in Friz and Victoir [8] may be a nice summary on this topic.

Throughout this paper, the time interval is $[0,1]$ and we set $\triangle = \{(s,t) \mid 0 \le s \le t \le 1\}$. For any $Y \in C(\triangle, \mathbf{R}^d)$, we set

$$\|Y\|_{\alpha-H} = \sup_{0 \le s < t \le 1} \frac{|Y_{s,t}|}{|t-s|^{\alpha}} \qquad (0 < \alpha \le 1), \tag{3.1}$$

$$\|Y\|_{\alpha,m-B} = \Big(\iint_{0 \le s < t \le 1} \frac{|Y_{s,t}|^m}{|t-s|^{1+m\alpha}} dsdt \Big)^{1/m} \qquad (m \ge 1,\ 0 < \alpha \le 1). \tag{3.2}$$

These are called α-Hölder norm and (α, m)-Besov norm, respectively. There are some variants of Besov-type norms, but we will use this one.

Set $C^{\alpha-H}([0,1], \mathbf{R}^d) = \{x \in C([0,1], \mathbf{R}^d) \mid \|X^1\|_{\alpha-H} < \infty\}$, where $X^1_{s,t} := x_t - x_s$. This is called the space of $\mathcal{V}$-valued α-Hölder continuous paths and becomes a real Banach space with $\|x\| = |x_0| + \|X^1\|_{\alpha-H}$. Its closed subset of paths that start at $a \in \mathbf{R}^d$ is denoted by $C^{\alpha-H}_a([0,1], \mathbf{R}^d)$. In a similar way, if $\alpha - m^{-1} > 0$, $C^{\alpha,m-B}([0,1], \mathbf{R}^d)$ etc. are defined. Then, $C^{m,\alpha-B}([0,1], \mathbf{R}^d)$ is continuously embedded in $C^{(\alpha-m^{-1})-H}([0,1], \mathbf{R}^d)$. (See Appendix A.2, [8].) In this case, closed subsets such as $C^{\alpha,m-B}_a([0,1], \mathbf{R}^d)$ are well-defined.

Let $T^2(\mathbf{R}^d) = \mathbf{R} \oplus \mathbf{R}^d \oplus (\mathbf{R}^d)^{\otimes 2}$ be the truncated tensor algebra of step 2. The set of elements of the form $(1, \bullet, \star)$ forms a non-abelian group under the tensor multiplication $\otimes$. The unit element is $\mathbf{1} = (1, 0, 0)$. Set

$$G^2(\mathbf{R}^d) = \{(1, a_1, a_2) \in T^2(\mathbf{R}^d) \mid a_2^{i,j} + a_2^{j,i} = a_1^i a_1^j \quad (1 \le i, j \le d)\}.$$

It is easy to check that $G^2(\mathbf{R}^d)$ becomes a subgroup. It is called the free nilpotent group of step 2. Note that the dilation on $T^2(\mathbf{R}^d)$ (i.e., $(1, a_1, a_2) \mapsto (1, \lambda a_1, \lambda^2 a_2)$ for $\lambda \in \mathbf{R}$) is well-defined on $G^2(\mathbf{R}^d)$, too.

A continuous map $X = (1, X^1, X^2) : \triangle \to T^2(\mathbf{R}^d)$ is called multiplicative if it satisfies that

$$X^1_{s,t} = X^1_{s,u} + X^1_{u,t}, \qquad X^2_{s,t} = X^2_{s,u} + X^2_{u,t} + X^1_{s,u} \otimes X^1_{u,t}, \qquad (s \le u \le t). \tag{3.3}$$

This relation is called Chen's identity and can also be written as $X_{s,t} = X_{s,u} \otimes X_{u,t}$. In particular, $X_{s,t}$ $(s \le t)$ is a "difference" of a group-valued path, since $X_{s,t} = (X_{0,s})^{-1} \otimes X_{0,t}$.

Let $1/3 < \alpha < 1/2$. The space of $\mathbf{R}^d$-valued α-Hölder rough path is defined by

$$\begin{aligned}\Omega^H_\alpha(\mathbf{R}^d) = \{X = (1, X^1, X^2) \in C(\triangle, T^2(\mathbf{R}^d)) \\ \mid \text{multiplicative and } \|X^1\|_{\alpha-H} < \infty, \|X^2\|_{2\alpha-H} < \infty\}.\end{aligned}$$

The topology of this space is naturally induced by the following distance: $d(X, Y) = \|X^1 - Y^1\|_{\alpha-H} + \|X^2 - Y^2\|_{2\alpha-H}$. In the same way, (α, m)-Besov rough path is defined for $m \ge 2$ and $1/3 < \alpha < 1/2$ with $\alpha - m^{-1} > 1/3$ as follows;

$$\begin{aligned}\Omega^B_{m,\alpha}(\mathbf{R}^d) = \{X = (1, X^1, X^2) \in C(\triangle, T^2(\mathbf{R}^d)) \\ \mid \text{multiplicative and } \|X^1\|_{\alpha,m-B} < \infty, \|X^2\|_{2\alpha,m/2-B} < \infty\}.\end{aligned}$$

The topology of this space is naturally induced by the following distance: $d(X, Y) = \|X^1 - Y^1\|_{\alpha,m-B} + \|X^2 - Y^2\|_{2\alpha,m/2-B}$. In what follows, we will often write $X = (X^1, X^2)$ for simplicity, since the 0th component "1" is obvious.

A Lipschitz continuous path (i.e., 1-Hölder continuous path) $x \in C_0^{1-H}([0, 1], \mathbf{R}^d)$ admits a natural lift to a rough path by setting

$$X^1_{s,t} := x_t - x_s, \qquad X^2_{s,t} := \int_s^t (x_u - x_s) \otimes dx_u, \qquad (s, t) \in \triangle.$$

It is easy to see that $X \in \Omega^H_\alpha(\mathbf{R}^d) \cap \Omega^B_{\alpha,m}(\mathbf{R}^d)$. We call a rough path X obtained in this way a smooth rough path lying above x, or the lift of x. The lift map is denoted by S_2, i.e., $X = S_2(x)$.

An α-Hölder weakly geometric rough path is $X \in \Omega^H_\alpha(\mathbf{R}^d)$ such that

$$X^{2;i,j}_{s,t} + X^{2;j,i}_{s,t} = X^{1;i}_{s,t} X^{1;j}_{s,t}, \qquad (1 \le i, j \le d, \quad (s, t) \in \triangle). \tag{3.4}$$

Here, $X_{s,t}^{2;i,j}$ stands for the (i,j)-component of $X_{s,t}^2$, etc. Obviously, a smooth rough path satisfies (3.4) and is an α-Hölder weakly geometric rough path. The set of α-Hölder weakly geometric rough paths is denoted by $G^w\Omega_\alpha^H(\mathbf{R}^d)$, which is a closed subset of $\Omega_\alpha^H(\mathbf{R}^d)$. In a similar way, $G^w\Omega_{\alpha,m}^B(\mathbf{R}^d)$ is defined. From (3.3) and (3.4), a weakly geometric rough path is a "difference" of a $G^2(\mathbf{R}^d)$-valued path.

Let $G\Omega_\alpha^H(\mathbf{R}^d)$ be the closure of the set of smooth rough paths, which is called the geometric rough path space with α-Hölder norm. The geometric rough path space $G\Omega_{\alpha,m}^B(\mathbf{R}^d)$ with (α,m)-Besov norm is similarly defined. Hence, we have the following inclusions;

$$G\Omega_\alpha^H(\mathbf{R}^d) \subset G^w\Omega_\alpha^H(\mathbf{R}^d) \subset \Omega_\alpha^H(\mathbf{R}^d), \qquad G\Omega_{\alpha,m}^B(\mathbf{R}^d) \subset G^w\Omega_{\alpha,m}^B(\mathbf{R}^d) \subset \Omega_{\alpha,m}^B(\mathbf{R}^d).$$

There is a natural correspondence between $G^2(\mathbf{R}^d)$-valued path spaces in the usual sense and weakly geometric rough path spaces. In order to explain that, we first introduce a distance d on $G^2(\mathbf{R}^d)$. Let $\|\mathbf{a}\|$ be the Carnot-Caratheodory norm as in Chapter 7, [8]. This is a homogeneous norm on $G^2(\mathbf{R}^d)$ with symmetry and subadditivity, too. Its explicit form is not needed in this paper. For our purpose, it is enough to keep in mind that, there exist a constant $c > 0$ such that

$$\frac{1}{c}\|\mathbf{a}\| \le |a_1|_{\mathbf{R}^d} + \sqrt{|a_2|_{\mathbf{R}^d\otimes\mathbf{R}^d}} \le c\|\mathbf{a}\| \qquad \text{for any } \mathbf{a} = (1, a_1, a_2) \in G^2(\mathbf{R}^d).$$

We set $d(\mathbf{a},\mathbf{b}) = \|\mathbf{a}^{-1} \otimes \mathbf{b}\|$. This defines a left-invariant distance on $G^2(\mathbf{R}^d)$, which induces the same topology as the relative one inherited from $T^2(\mathbf{R}^d)$.

If $\sup_{0\le s<t\le 1} d(g_s, g_t)|t-s|^{-\alpha} < \infty$, a $G^2(\mathbf{R}^d)$-valued continuous path g is said to be α-Hölder continuous. The set of such paths starting at $\mathbf{1}$ is denoted by $C_{\mathbf{1}}^{\alpha-H}([0,1], G^2(\mathbf{R}^d))$. Similarly, g is said to be (α,m)-Besov if $\iint_{0\le s<t\le 1} d(g_s,g_t)^m |t-s|^{-(1+m\alpha)} ds dt < \infty$. The set of such paths starting at $\mathbf{1}$ is denoted by $C_{\mathbf{1}}^{\alpha,m-B}([0,1], G^2(\mathbf{R}^d))$. For a weakly geometric rough path X, we can associate a $G^2(\mathbf{R}^d)$-valued continuous path $t \mapsto X_{0,t}$. This defines a natural bijection between $C_{\mathbf{1}}^{\alpha-H}([0,1], G^2(\mathbf{R}^d))$ and $G^w\Omega_\alpha^H(\mathbf{R}^d)$. In the same way, there is a natural bijection between $C_{\mathbf{1}}^{\alpha,m-B}([0,1], G^2(\mathbf{R}^d))$ and $G^w\Omega_{m,\alpha}^B(\mathbf{R}^d)$. Through these bijections, we introduce distance functions on $C_{\mathbf{1}}^{\alpha-H}([0,1], G^2(\mathbf{R}^d))$ and $C_{\mathbf{1}}^{m,\alpha-B}([0,1], G^2(\mathbf{R}^d))$, respectively.

Proposition 3.1 *Assume* $1/3 < \alpha < 1/2$, $m \ge 2$, *and* $\alpha - 1/m > 1/3$. *Then,* $G\Omega_{\alpha,m}^B(\mathbf{R}^d)$ *is continuously embedded in* $G\Omega_{\alpha-1/m}^H(\mathbf{R}^d)$.

Proof. It is sufficient to show that $W\Omega_{\alpha,m}^B(\mathbf{R}^d) \cong C_{\mathbf{1}}^{\alpha,m-B}([0,1], G^2(\mathbf{R}^d))$ is continuously embedded in $W\Omega_{\alpha-1/m}^H(\mathbf{R}^d) \cong C_{\mathbf{1}}^{(\alpha-m^{-1})-H}([0,1], G^2(\mathbf{R}^d))$. The inclusion, which is called Besov-Hölder embedding, is shown in Corollary A.2, [8]. Its continuity is shown in Proposition A.9, [8]. ∎

Proposition 3.2 *Assume $1/3 < \alpha < \alpha' < 1/2$, $m \ge 2$, and $\alpha - 1/m > 1/3$. Then, the injection $G\Omega^B_{\alpha',m}(\mathbf{R}^d) \hookrightarrow G\Omega^B_{\alpha,m}(\mathbf{R}^d)$ maps a bounded subset to a precompact subset.*

Proof. Let $\{X(n)\}_{n=1}^{\infty}$ be any bounded sequence in $G\Omega^B_{\alpha',m}(\mathbf{R}^d)$. Define $g(n)_t = X(n)_{0,t}$. Then, by Proposition 3.1, $\{g(n)\}_{n=1}^{\infty}$ is a bounded sequence in $C_{\mathbf{1}}^{(\alpha-m^{-1})-H}([0,1], G^2(\mathbf{R}^d))$. Hence, as functions of t, $\{g(n)\}_{n=1}^{\infty}$ is uniformly bounded and equicontinuous. By Ascoli-Arzela's theorem, there exists a subsequence, which is denoted by $\{g(n)\}$ again, which converges to $g \in C_{\mathbf{1}}([0,1], G^2(\mathbf{R}^d))$ in sup-norm. Set $X_{s,t} = g_s^{-1} \otimes g_t$ for all $(s,t) \in \triangle$. It is sufficient to show that $\lim_{n\to\infty} X(n) = X$ in (α, m)-Besov topology. First, note that $\lim_{n\to\infty} \sup_{(s,t)\in\triangle} |X(n)^j_{s,t} - X^j_{s,t}| = 0$ for $j = 1, 2$. Set $r = (1+m\alpha')/(1+m\alpha) > 1$ and let r' be its conjugate exponent. Then, by Hölder's inequality,

$$
\begin{aligned}
&\iint_{0\le s<t\le 1} \frac{|X(n)^1_{s,t} - X(n')^1_{s,t}|^m}{|t-s|^{1+m\alpha}} dsdt \\
&\le \sup_{(s,t)\in\triangle} |X(n)^1_{s,t} - X(n')^1_{s,t}|^{m(1-1/r)} \iint_{0\le s<t\le 1} \frac{|X(n)^1_{s,t} - X(n')^1_{s,t}|^{m/r}}{|t-s|^{1+m\alpha}} dsdt \\
&\le \sup_{(s,t)\in\triangle} |X(n)^1_{s,t} - X(n')^1_{s,t}|^{m(1-1/r)} \Bigl(\iint_{0\le s<t\le 1} \frac{|X(n)^1_{s,t} - X(n')^1_{s,t}|^m}{|t-s|^{1+m\alpha'}} dsdt \Bigr)^{1/r} \frac{1}{2^{1/r'}} \\
&\le \text{const.} \times \sup_{(s,t)\in\triangle} |X(n)^1_{s,t} - X(n')^1_{s,t}|^{m(1-1/r)} \to 0 \qquad \text{as } n, n' \to \infty.
\end{aligned}
$$

Here, we used that (α', m)-Besov norm of $X(n)$ are bounded. The limit in the Besov sense must coincide with X^1. The second level path can be dealt with in the same way. ∎

4 Preliminaries from quasi-sure analysis

In this section we recall basics of Malliavin calculus and, in particular, of quasi-sure analysis. Generalized Wiener functionals (i.e., Watanabe distributions) and capacities associated with Gaussian Sobolev spaces play important roles in this paper. Analysis of these objects is called quasi-sure analysis.

4.1 Basics of Malliavin calculus

In this subsection, we recall the basic notions in Malliavin calculus, in particular, the theory of Gaussian Sobolev spaces. We mainly follow Sections 5.8 –5.10, [11], or Shigekawa [21].

Let $(\mathcal{W}, \mathcal{H}, \mu)$ be an abstract Wiener space. As usual, $\mathcal{H}$ and $\mathcal{H}^*$ are identified through Riesz isometry. Let D be the $\mathcal{H}$-derivative and D^* be its dual. Ornstein-Uhlenbeck operator is denoted by $L = -D^*D$. The first Wiener chaos associated with $h \in \mathcal{H}$

is denoted by $\langle h, w\rangle$. If $\langle h, \cdot\rangle \in \mathcal{W}^*$, then it coincides with ${}_{\mathcal{W}^*}\langle h, w\rangle_{\mathcal{W}}$. A function $F : \mathcal{W} \to \mathbf{R}$ is called a real-valued polynomial if there exist $m \in \mathbf{N}$, $h_1, \ldots, h_m$, and a real-valued polynomial f of m-variables such that $F(w) = f(\langle h_1, w\rangle, \ldots, \langle h_m, w\rangle)$. The set of all real-valued polynomials are denoted by $\mathcal{P}$. Let $\mathcal{K}$ be a real separable Hilbert space. $L^q(\mathcal{K}) = L^q(\mathcal{W}; \mathcal{K})$ denotes the L^q-space of $\mathcal{K}$-valued functions. A function $G : \mathcal{W} \to \mathcal{K}$ is called a $\mathcal{K}$-valued polynomial if there exist $m \in \mathbf{N}_+$, $F_1, \ldots, F_m \in \mathcal{P}$, and $v_1, \ldots, v_m \in \mathcal{K}$ such that $G(w) = \sum_{j=1}^m F_j(w)v_j$. The set of all $\mathcal{K}$-valued polynomials are denoted by $\mathcal{P}(\mathcal{K})$.

For $q \in (1, \infty)$ and $r \in \mathbf{R}$, and $F \in \mathcal{P}(\mathcal{K})$, set $\|F\|_{q,r} = \|(I - L)^{-r/2}F\|_q$. We define the Sobolev space $\mathbf{D}_{q,r}(\mathcal{K})$ to be the completion of $\mathcal{P}(\mathcal{K})$ with respect to this norm. When $\mathcal{K} = \mathbf{R}$, we simply write L^q, $\mathbf{D}_{q,r}$, etc. If $q \le q'$ and $r \le r'$, then $\|\cdot\|_{q,r} \le \|\cdot\|_{q',r'}$ and $\mathbf{D}_{q,r}(\mathcal{K}) \supset \mathbf{D}_{q',r'}(\mathcal{K})$.

When $r = k \in \mathbf{N}$, then the n Meyer equivalence holds (see [21] for example); there exists a positive constant $c_{q,r}$ such that

$$c_{q,k}^{-1}\|F\|_{q,k} \le \|F\|_q + \|D^kF\|_q \le c_{q,k}\|F\|_{q,k} \qquad \text{for all } F \in \mathbf{D}_{q,k}(\mathcal{K}).$$

Now we discuss the Wiener chaos. Set $\hat{\mathcal{C}}_n$ be the L^2-closure of real-valued polynomials of order $\le n$. In particular, $\hat{\mathcal{C}}_0$ is the space of constant functions on $\mathcal{W}$. $\hat{\mathcal{C}}_n$ is called the inhomogeneous Wiener chaos of order n. Set $\mathcal{C}_0 = \hat{\mathcal{C}}_0$ and $\mathcal{C}_n = \hat{\mathcal{C}}_n \cap \mathcal{C}_{n-1}^{\perp}$. This is called the homogeneous Wiener chaos of order n. Note that $\mathcal{C}_1 = \{\langle h, \cdot\rangle \mid h \in \mathcal{H}\}$. It is known that $\mathcal{C}_n$ is the eigenspace of $-L$ in L^2 that corresponds to the eigenvalue n. Clearly, $\hat{\mathcal{C}}_n = \oplus_{k=0}^n \mathcal{C}_k$.

Proposition 4.1 *For any $n \in \mathbf{N}$, $q \in (1, \infty)$ and $r \ge 0$, there exists a constant $M_{n,q,r} \ge 1$ such that*

$$M_{n,q,r}^{-1}\|F\|_2 \le \|F\|_{q,r} \le M_{n,q,r}\|F\|_2 \qquad (F \in \hat{\mathcal{C}}_n). \tag{4.1}$$

In other words, restricted on $\hat{\mathcal{C}}_n$, all (q, r)-Sobolev norms are equivalent.

Proof. It is sufficient to show (4.1) for $\mathcal{C}_n$, instead of $\hat{\mathcal{C}}_n$. It is shown in Proposition 2.14, [21] that all L^q-norms are equivalent on $\mathcal{C}_n$. Using the fact that $\mathcal{C}_n$ is an eigenspace of $-L$, we can easily see that all (q, r)-Sobolev norms are equivalent on $\mathcal{C}_n$. ∎

Now we recall generalized Wiener functionals, which are also called Watanabe distributions. Set $\mathbf{D}_\infty(\mathcal{K}) = \cap_{1<q<\infty, r\in\mathbf{R}}\mathbf{D}_{q,r}(\mathcal{K})$ and $\mathbf{D}_{-\infty}(\mathcal{K}) = \cup_{1<q<\infty, r\in\mathbf{R}}\mathbf{D}_{q,r}(\mathcal{K})$. Those are called the space of test functions and the space of generalized Wiener functionals, respectively. We also use the following spaces of (generalized) Wiener functionals: $\tilde{\mathbf{D}}_\infty(\mathcal{K}) = \cap_{k=1}^\infty \cup_{1<q<\infty} \mathbf{D}_{q,k}(\mathcal{K})$ and $\tilde{\mathbf{D}}_{-\infty}(\mathcal{K}) = \cup_{k=1}^\infty \cap_{1<q<\infty} \mathbf{D}_{q,-k}(\mathcal{K})$. In other words, $F \in \tilde{\mathbf{D}}_\infty(\mathcal{K})$ is equivalent to that, for any $k \ge 0$, there exists $q = q(k) > 1$ such that $F \in \mathbf{D}_{q,k}(\mathcal{K})$.

Let $F = (F^1, \ldots, F^n) \in \mathbf{D}_\infty(\mathbf{R}^n)$. The Malliavin covariance matrix is defined by $\big(\langle DF^i(w), DF^j(w)\rangle_{\mathcal{H}}\big)_{1\le i,j\le n}$. We say that F is non-degenerate in the sense of Malliavin if $\det\big(\langle DF^i(w), DF^j(w)\rangle_{\mathcal{H}}\big)^{-1}_{1\le i,j\le n}$ is in L^q for all $1 < q < \infty$. This non-degeneracy is very important in Malliavin calculus. For example, such a non-degenerate F can be composed with a Schwartz distribution ϕ defined on $\mathbf{R}^n$ and $\phi \circ F$ becomes a generalized Wiener functional.

We introduce seminorms on the space of tempered distributions on $\mathbf{R}^n$. For $k \in \mathbf{Z}$ and a real-valued, rapidly decreasing, smooth function ϕ of Schwartz class $\mathcal{S}(\mathbf{R}^n)$ on $\mathbf{R}^n$, we define $\|\phi\|_{2k} = \|(1 + |\cdot|^2 - \triangle/2)^k \phi\|_\infty$. Set $\mathcal{S}_{2k}(\mathbf{R}^n)$ to be the completion of $\mathcal{S}(\mathbf{R}^n)$ with respect to this norm. Then $\mathcal{S}(\mathbf{R}^n) = \cap_{k>0}\mathcal{S}_{2k}(\mathbf{R}^n)$, which is a Fréchet space. The dual space $\mathcal{S}'(\mathbf{R}^n) = \cup_{k>0}\mathcal{S}_{-2k}(\mathbf{R}^n)$ is called the Schwartz space of tempered distributions.

S. Watanabe proved that, if a Wiener functional F is non-degenerate, then the pullback map $\phi \mapsto \phi \circ F$ extends to a continuous linear map between two distribution spaces. The following is borrowed from pp. 378-379, Ikeda and Watanabe [11].

Theorem 4.2 *Let $F = (F^1, \ldots, F^n) \in \mathbf{D}_\infty(\mathbf{R}^n)$ be non-degenerate in the sense of Malliavin. Then, for any $1 < q < \infty$ and $k = 0, 1, 2, \ldots$, there exists a positive constant $C = C(q, k, F)$ such that*

$$\|\phi \circ F\|_{q,-2k} \le C\|\phi\|_{-2k} \qquad (\phi \in \mathcal{S}(\mathbf{R}^n))$$

holds. Therefore, the map $\phi \mapsto \phi \circ F$ extends uniquely to a continuous linear map $\mathcal{S}_{-2k}(\mathbf{R}^n) \ni T \mapsto T \circ F \in \mathbf{D}_{q,-2k}$. In particular, $T \circ F \in \cup_{k=1}^\infty \cap_{1<q<\infty} \mathbf{D}_{q,-2k} = \tilde{\mathbf{D}}_{-\infty}$.

We recall an integration by parts formula in the context of Malliavin calculus (See p. 377, [11]). For $F = (F^1, \ldots, F^n) \in \mathbf{D}_\infty(\mathbf{R}^n)$, we denote by $\tau^{ij}(w) = \langle DF^i(w), DF^j(w)\rangle_{\mathcal{H}}$ the (i,j)-component of Malliavin covariance matrix. We denote by $\gamma^{ij}(w)$ the (i,j)-component of the inverse matrix τ^{-1}. Note that $\tau^{ij} \in \mathbf{D}_\infty$ and $D\gamma^{ij} = \sum_{k,l}\gamma^{ik}(D\tau^{kl})\gamma^{lj}$. Hence, derivatives of γ^{ij} can be written in terms of γ^{ij}'s and the derivatives of τ^{ij}'s. Suppose $G \in \mathbf{D}_\infty$ and $T \in \mathcal{S}'(\mathbf{R}^n)$. Then, the following integration by parts holds;

$$\mathbb{E}\big[(\partial_i T \circ F) \cdot G\big] = \mathbb{E}\big[(T \circ F) \cdot \Phi_i(\,\cdot\,; G)\big] \tag{4.2}$$

where $\Phi_i(\,\cdot\,; G) \in \mathbf{D}_\infty$ is given by

$$\begin{aligned}\Phi_i(w; G) = -\sum_{j=1}^d \Big\{ -\sum_{k,l=1}^d G(w)\gamma^{ik}(w)\gamma^{jl}(w)\langle D\tau^{kl}(w), DF^j(w)\rangle_{\mathcal{H}} \\ + \gamma^{ij}(w)\langle DG(w), DF^j(w)\rangle_{\mathcal{H}} + \gamma^{ij}(w)G(w)LF^j(w)\Big\}.\end{aligned} \tag{4.3}$$

Note that the expectations in (4.2) are in fact the pairing of $\tilde{\mathbf{D}}_{-\infty}$ and $\tilde{\mathbf{D}}_\infty$.

4.2 Basics of capacities

In this subsection, we recall the definition and basic properties of the capacity associated with the Sobolev space $\mathbf{D}_{q,r}$ for $1 < q < \infty$ and $r \in \mathbf{N}$.

The contents of the subsection is borrowed from Chapter 9, Malliavin [17]. In this book, they work on a particular Gaussian space, namely $\mathbf{R}^\infty$ equipped with countable product of one-dimensional standard normal distribution. But, the results in [17], at least the ones we will use in this paper, hold true on any abstract Wiener space. In some literatures, a slightly different definition of capacities is used (see [22] for instance).

Let $(\mathcal{W}, \mathcal{H}, \mu)$ be an abstract Wiener space. We keep the same notation as in the previous subsection. Throughout this subsection, we set $1 < q < \infty$ and $r \in \mathbf{N}$. For an open subset $O \subset \mathcal{W}$, we set

$$\mathrm{c}_{q,r}(O) = \inf\{\|\phi\|_{q,r} \mid \phi(w) \geq 1 \text{ a.e. on } O \text{ and } \phi(w) \geq 0 \text{ a.e. on } \mathcal{W}\}. \tag{4.4}$$

For any subset $A \subset \mathcal{W}$, which is not necessarily open, we define

$$\mathrm{c}_{q,r}(A) = \inf\{\mathrm{c}_{q,r}(O) \mid A \subset O \text{ and } O \text{ is open }\}. \tag{4.5}$$

Since (q, r)-norm is increasing in both q and r, so is $c_{q,r}(A)$. We say that a property $\pi = \pi_w$ holds (q, r)-quasi-everywhere if $\mathrm{c}_{q,r}(\{w | \pi_w \text{ does not hold}\}) = 0$. We say that a property $\pi = \pi_w$ holds quasi-surely if it holds (q, r)-quasi-everywhere for all $q \in (1, \infty)$ and $r \in \mathbf{N}$. A subset A is called slim if $\mathrm{c}_{q,r}(A) = 0$ for all $q \in (1, \infty)$ and $r \in \mathbf{N}$.

Now let us discuss quasi-continuity. A function ϕ from $\mathcal{W}$ to $\mathbf{R}$ (or to a metric space) is said to be (q, r)-quasi-continuous if, for any $\varepsilon > 0$, there exists an open set O_ε such that $\mathrm{c}_{q,r}(O_\varepsilon) < \varepsilon$ and the restriction $\phi|_{O_\varepsilon^c}$ is continuous. A function ϕ from $\mathcal{W}$ to $\mathbf{R}$ (or to a metric space) is said to be ∞-quasi-continuous if it is (q, r)-quasi-continuous for all q and r. This is equivalent to that there is a decreasing sequence of open subsets O_n such that $\lim_{n\to\infty} \mathrm{c}_{n,n}(O_n) = 0$ and $\phi|_{O_n^c}$ is continuous. (In this paragraph, the set O_ε or O_n actually need not be open, since any subset can be approximated in terms of capacity by an open subset from outside.)

For a measurable function ψ, we say ψ^* is a (q, r)-redefinition of ψ if ψ^* is (q, r)-quasi-continuous and $\psi = \psi^*$ a.s. (ψ^* is also called (q, r)-quasi continuous modification.) Note that (q, r)-redefinition is essentially unique when it exists. It is shown in Theorem 2.3.3, [17] that any $\phi \in \mathbf{D}_{q,r}$ admits a (q, r)-redefinition. Similarly, for a measurable function ψ, we say ψ^* is a ∞-redefinition of ψ if ψ^* is ∞-quasi-continuous and $\psi = \psi^*$ a.s. It is shown in Subsection 2.4, [17] that any $\phi \in \mathbf{D}_\infty$ admits an (essentially unique) ∞-redefinition.

Now we give two useful basic lemmas without proofs for later use. One is Borel-Cantelli's lemma (Corollary 1.2.4, [17]) and the other is Chebyshev's lemma (Theorem 2.2, [17]).

Lemma 4.3 *Let $1 < q < \infty$ and $r \in \mathbf{N}$. Assume that $A_k \subset \mathcal{W}$ $(k \in \mathbf{N})$ satisfy that $\sum_k \mathrm{c}_{q,r}(A_k) < \infty$. Then, $\mathrm{c}_{q,r}(\limsup_{k\to\infty} A_k) = 0$.*

Lemma 4.4 *For any $1 < q < \infty$ and $r \in \mathbf{N}$, there exists a positive constant $M_{q,r}$ such that, for any $\phi \in \mathbf{D}_{q,r}$ and any $R > 0$, we have*

$$\mathrm{c}_{q,r}(\{w \mid \phi^*(w) > R\}) \le R^{-1} M_{q,r} \|\phi\|_{q,r}.$$

In the finite dimensional calculus, it is n that a positive Schwartz distribution is a measure. An analogous fact is true in Malliavin calculus, too. It is called Sugita's theorem (Theorem 3.0, [17] or Sugita [22]) and will play an important role in the sequel.

Proposition 4.5 *Let $l \in \mathbf{D}_{-\infty}$ be a positive Watanabe distribution, that is, it satisfies ${}_{\mathbf{D}_{-\infty}}\langle l, f \rangle_{\mathbf{D}_\infty} = \mathbb{E}[l \cdot f] \ge 0$ for any $f \in \mathbf{D}_\infty$ such that $f > 0$ a.s. Then, there exists a unique positive Borel measure θ of finite total mass such that*

$$\langle l, g \rangle = \mathbb{E}[l \cdot g] = \int_{\mathcal{W}} g^*(w) d\theta(w), \qquad (g \in \mathbf{D}_\infty). \tag{4.6}$$

Furthermore, θ does not charge a slim set. (Hence, any choice of ∞-redefinition g^* will do.)

It is easy to see that, if $l \in \mathbf{D}_{q',-r}$ with $1/q + 1/q' = 1$, then $\theta(O) \le \mathrm{c}_{q,r}(O)\|l\|_{q',-r}$ for any open set O. (Note that O need not be open here again.)

Let us discuss the equilibrium potential. Given a Borel subset $A \subset \mathcal{W}$, we define

$$\mathcal{F}_{q,r}(A) = \{u \in \mathbf{D}_{q,r} \mid \ u^* \ge 1 \ (q,r)\text{-quasi-everywhere on } A \ \}. \tag{4.7}$$

This is a closed convex subset of $\mathbf{D}_{q,r}$ and has a unique element $\phi_A \in \mathcal{F}_{q,r}(A)$, which minimizes (q,r)-Sobolv norm. We call ϕ_A the equilibrium potential of A. Then, Theorem 4.4, [17] states that,

$$\mathrm{c}_{q,r}(A) = \|\phi_A\|_{q,r} \le \|u\|_{q,r} \qquad (u \in \mathcal{F}_{q,r}(A)). \tag{4.8}$$

5 Quasi-sure existence of Brownian rough path

In this section, we recall that Brownian motion admits a natural lift quasi-surely via the dyadic partitions. This fact was proved by three authors independently, Higuchi [9], Inahama [12], and Watanabe [25]. Among them, Higuchi's method seems best. Higuchi's master thesis is in Japanense and probably unavailable outside Japan. However, Section 3 of Aida's recent paper [1] is essentially the same. So we will follow [9, 1], in which a slightly different Besov norm is used.

From now on, we denote by $(\mathcal{W}, \mathcal{H}, \mu)$ be the d-dimensional classical Wiener space. That is, $\mathcal{W} = C_0([0,1], \mathbf{R}^d)$ with the sup-norm, $\mathcal{H}$ is the Cameron-Martin space, and μ is the usual d-dimensional Wiener measure. For $w \in \mathcal{W}$ and $k \in \mathbf{N}$, $w(k) \in C_0^{1-H}([0,1], \mathbf{R}^d)$ denotes the mth dyadic polygonal approximation associated with the partition $\{l 2^{-k} \mid 0 \le l \le 2^k\}$ of $[0,1]$. We denote by $W(k) := S_2(w(k))$ the natural lift of $w(k)$.

For α, m such that $4m \geq 2$, $1/3 < \alpha < 1/2$, and $\alpha - 1/(4m) > 1/3$, we set

$$\mathcal{Z}_{\alpha,4m} = \{ w \in \mathcal{W} \mid \ \{W(k)\}_{k=1}^{\infty} \text{ is Cauchy in } G\Omega^B_{\alpha,4m}(\mathbf{R}^d) \}.$$

Slightly abusing the notation, we denote $\lim_{k\to\infty} W(k)$ by W or $S_2(w)$ for $w \in \mathcal{Z}_{\alpha,4m}$. If $w \notin \mathcal{Z}_{\alpha,4m}$, $S_2(w)$ is not defined. (So, as a map from $\mathcal{W}$, S_2's definition depends on α, m.) The aim of this section is to prove that, for sufficiently large $m \in \mathbf{N}$, $\mathcal{Z}^c_{\alpha,4m}$ is a slim set.

Some basic properties of $\mathcal{Z}_{\alpha,4m}$ are shown in the following lemma. We can see from this that we may write cW ($c \in \mathbf{R}$) without ambiguity.

Lemma 5.1 *Let α, m and $\mathcal{Z}_{\alpha,4m}$ be as above. Then, we have the following* (i)–(iii)*.*
(i) $\mathcal{H} \subset \mathcal{Z}_{\alpha,4m}$.
(ii) *For any $c \in \mathbf{R}$ and $w \in \mathcal{Z}_{\alpha,4m}$, $S_2(cw) = cS_2(w)$. In particular, $c\mathcal{Z}_{\alpha,4m} = \mathcal{Z}_{\alpha,4m}$ if $c \neq 0$.*
(iii) *Assume in addition that $2m - 4m\alpha - 1 = 4m(1/2 - \alpha) - 1 > 0$. Then, for any $h \in \mathcal{H}$ and $w \in \mathcal{Z}_{\alpha,4m}$, $S_2(w + h) = T_h(S_2(w))$. Here, the right hand side stands for the Young translation of $S_2(w)$ by h. In particular, $T_h(\mathcal{Z}_{\alpha,4m}) = \mathcal{Z}_{\alpha,4m}$*

Proof. Throughout this proof, the constant $C > 0$ may change from line to line. (ii) is trivial. Note that $\|h\|_{1/2-H} \leq \|h\|_{\mathcal{H}}$ and that, for any h, $\|h(k) \to h\|_{\mathcal{H}} \to 0$ as $k \to \infty$. Then, (i) is immediate from these.

We now prove (iii). The only non-trivial part is boundedness of the "cross terms" of the second level path as bilinear functionals in X^1 and h. For $x \in C_0^{\alpha,4m-B}([0,1], \mathbf{R}^d)$ and $h \in \mathcal{H}$, consider

$$J[x,h]_{s,t} := \int_s^t (x_u - x_s) \otimes dh_u.$$

Then, we can easily see from (3.1) that

$$\begin{aligned} |J[x,h]_{s,t}|^2 &\leq \Big(\int_s^t |x_u - x_s||h'_u| du \Big)^2 \leq \|h\|^2_{\mathcal{H}} \int_s^t |x_u - x_s|^2 du \\ &\leq C\|h\|^2_{\mathcal{H}} \|x\|^2_{\alpha,4m-B} \int_s^t (u-s)^{2(\alpha - 1/4m)} du \\ &\leq C\|h\|^2_{\mathcal{H}} \|x\|^2_{\alpha,4m-B} (t-s)^{2(\alpha-1/4m)+1} \end{aligned}$$

Then, we have

$$\begin{aligned} \|J[x,h]\|^{2m}_{2\alpha,2m-B} &\leq C\|h\|^{2m}_{\mathcal{H}} \|x\|^{2m}_{\alpha,4m-B} \iint_{0\leq s\leq t\leq 1} \frac{(t-s)^{2m(\alpha-1/4m)+m}}{(t-s)^{1+4m\alpha}} dsdt \\ &\leq C\|h\|^{2m}_{\mathcal{H}} \|x\|^{2m}_{\alpha,4m-B} \iint_{0\leq s\leq t\leq 1} \frac{1}{(t-s)^{1+2m\alpha-m+1/2}} dsdt \\ &\leq C\|h\|^{2m}_{\mathcal{H}} \|x\|^{2m}_{\alpha,4m-B}. \end{aligned}$$

In the same way, $J[h,x]$ can be estimated. The rest is a routine and we omit it. ∎

Take any $\varepsilon > 0$ and fix it. We set

$$\mathcal{A}^k_{\alpha,4m} = \{w \in \mathcal{W} \mid \|w(k+1) - w(k)\|^{4m}_{\alpha,4m-B} \ge k^{-(4m+\varepsilon)}\}, \tag{5.1}$$

Note that $\|w(k+1) - w(k)\|^{4m}_{\alpha,4m-B}$ is in the inhomogeneous Wiener chaos $\hat{\mathcal{C}}_{4m}$ if $m \in \mathbf{N}$. We can estimate the capacities of $\mathcal{A}^k_{\alpha,4m}$ as follows.

Lemma 5.2 *As before, let $m \in \mathbf{N}_+$, $1/3 < \alpha < 1/2$, and $\alpha-(1/4m) > 1/3$. Furthermore, we assume that $4m - 8m\alpha - 1 = 8m(1/2-\alpha) - 1 > 0$. Then, for any $1 < q < \infty$ and $r \in \mathbf{N}$,*

$$\sum_{k=1}^{\infty} \mathrm{c}_{q,r}(\mathcal{A}^k_{\alpha,4m}) \le M_{q,r} \sum_{k=1}^{\infty} k^{4m+\varepsilon} \Big\| \|w(k+1) - w(k)\|^{4m}_{\alpha,4m-B} \Big\|_{q,r} < \infty. \tag{5.2}$$

Here, $M_{q,r} > 0$ is the constant given in Lemma 4.4. As a result, $\limsup_{k\to\infty} \mathcal{A}^k_{\alpha,4m}$ is slim, which implies quasi-sure convergence of $\{W(k)^1\}_k$.

Proof. The left inequality in (5.2) is immediate from Lemma 4.4. First we will prove the summability for L^2-norm, instead of $\mathbf{D}_{q,r}$-norm.

We set $z(k) = w(k+1) - w(k)$ and $Z(k)^1_{s,t} = z(k)_t - z(k)_s$ for simplicity. It can be written explicitly as follows;

$$\begin{aligned} z(k)_t = 2^k \{ (t - \frac{j-1}{2^k}) \wedge (\frac{j}{2^k} - t) \} \\ \times \Big(2w(\frac{2j-1}{2^{k+1}}) - w(\frac{j-1}{2^k}) - w(\frac{j}{2^k}) \Big) \qquad \text{if } \frac{j-1}{2^k} \le t \le \frac{j}{2^k}. \end{aligned} \tag{5.3}$$

Note that $\|z(k)\|^{4m}_{\alpha,4m-B} \in \hat{\mathcal{C}}_{4m}$. A straight forward computation yields;

$$\mathbb{E}[|Z(k)^1_{s,t}|^2] \le C_d \big(\frac{1}{2^k} \wedge (t-s) \big), \qquad (s,t) \in \triangle \tag{5.4}$$

for some constant $C_d > 0$, which depends only on d. Then, we have

$$\begin{aligned} \Big\| \|z(k)\|^{4m}_{\alpha,4m-B} \Big\|_2^2 &\le \mathbb{E}\Big[\Big(\iint_{0\le s\le t\le 1} \frac{|Z(k)^1_{s,t}|^{4m}}{(t-s)^{1+4m\alpha}} dsdt \Big)^2 \Big] \\ &\le \frac{1}{2} \iint_{0\le s\le t\le 1} \frac{\mathbb{E}[|Z(k)^1_{s,t}|^{8m}]}{(t-s)^{2+8m\alpha}} dsdt \\ &\le C_{m,d} \iint_{0\le s\le t\le 1} \frac{(2^{-k})^{4m} \wedge (t-s)^{4m}}{(t-s)^{2+8m\alpha}} dsdt. \end{aligned} \tag{5.5}$$

Here, we used Schwarz's inequality and Proposition 4.1 for $Z(k)^1_{s,t} \in \mathcal{C}_1$.

Now use the following well-known inequality;

$$\iint_{0\le s\le t\le 1} \frac{\eta \wedge (t-s)^a}{(t-s)^b} dsdt \le \frac{a\eta^{(a-b+1)/a}}{(a-b+1)(b-1)} \qquad (0 < \eta \le 1, b > 1, a > b-1). \tag{5.6}$$

To check this formula, just change variables by $u = s$, $v = t - s$. Then, the domain of integral becomes $\{0 < u < 1, 0 < v < 1, u + v < 1\}$. The rest is easy.

Observe that, if $4m - 8m\alpha - 1 = 8m(1/2 - \alpha) - 1 > 0$, we can use (5.6). We have

$$\Big\| \|z(k)\|_{\alpha,4m-B}^{4m} \Big\|_2 \le C_{\alpha,m,d} \Big(\frac{1}{2^{(4m-8m\alpha-1)/2}} \Big)^k.$$

This proves the summability for L^2-norm. As for $\mathbf{D}_{q,r}$-norm, we can easily see from Proposition 4.1 that

$$\begin{aligned} \sum_{k=1}^{\infty} k^{4m+\varepsilon} \Big\| \|z(k)\|_{\alpha,4m-B}^{4m} \Big\|_{q,r} &\le M_{4m,q,r} \sum_{k=1}^{\infty} k^{4m+\varepsilon} \Big\| \|z(k)\|_{\alpha,4m-B}^{4m} \Big\|_2 \\ &\le C_{q,r,\alpha,m,d} \sum_{k=1}^{\infty} k^{4m+\varepsilon} \Big(\frac{1}{2^{(4m-8m\alpha-1)/2}} \Big)^k < \infty. \end{aligned} \tag{5.7}$$

Here, $c_{q,r,\alpha,m,d}$ is a positive constant which depends only on q, r, α, m, d.

It is immediate from Lemma 4.3 that $\limsup_{k\to\infty} \mathcal{A}_{\alpha,4m}^k$ is slim. Note that $w \in (\limsup_{k\to\infty} \mathcal{A}_{\alpha,4m}^k)^c = \cup_{N=1}^{\infty} \cap_{k=N}^{\infty} (\mathcal{A}_{\alpha,4m}^k)^c$ implies that $\{w(k)\} = \{W(k)^1\}$ is convergent in (α, m)-Besov norm. ∎

Next, let us consider the second level paths. For $\mathbf{R}^d$-valued continuous paths x and y, we define $J[x, y] : \triangle \to (\mathbf{R}^d)^{\otimes 2}$ by

$$J[x, y]_{s,t} = \int_s^t (x_u - x_s) \otimes dy_u,$$

whenever the integral on the right hand side can be defined. Note that the following equality holds;

$$\begin{aligned} J[x, x]_{s,t} - J[y, y]_{s,t} = J[x - y, x - y]_{s,t} &+ J[x - y, y]_{s,t} \\ &- J[x - y, y]_{s,t}^* + Y_{s,t}^1 \otimes (X_{s,t}^1 - Y_{s,t}^1). \end{aligned} \tag{5.8}$$

Here, $*$ stands for the linear isometry of $(\mathbf{R}^d)^{\otimes 2}$ defined by $(\xi \otimes \eta)^* = \eta \otimes \xi$ for all $\xi, \eta \in \mathbf{R}^d$. Indeed, we can easily see that

$$J[x - y, x - y]_{s,t} - J[x, x]_{s,t} + J[y, y]_{s,t} = -J[x - y, y]_{s,t} - J[y, x - y]_{s,t}$$

and, from integration by parts, that

$$J[y, x - y]_{s,t} = Y_{s,t}^1 \otimes (X_{s,t}^1 - Y_{s,t}^1) - \int_s^t dy_u \otimes (X_{s,u}^1 - Y_{s,u}^1).$$

Thus, we have verified (5.8).

With (5.8) in hand, we naturally consider the following subsets in $\mathcal{W}$.

$$\mathcal{B}(1)^k_{2\alpha,2m} = \{w \in \mathcal{W} \mid \|J[z(k), z(k)]\|^{2m}_{2\alpha,2m-B} \geq k^{-(2m+\varepsilon)}\}, \tag{5.9}$$
$$\mathcal{B}(2)^k_{2\alpha,2m} = \{w \in \mathcal{W} \mid \|J[z(k), w(k)]\|^{2m}_{2\alpha,2m-B} \geq k^{-(2m+\varepsilon)}\}, \tag{5.10}$$
$$\mathcal{B}(3)^k_{2\alpha,2m} = \{w \in \mathcal{W} \mid \|W(k)^1 \otimes Z(k)^1\|^{2m}_{2\alpha,2m-B} \geq k^{-(2m+\varepsilon)}\}, \tag{5.11}$$

where we set $z(k) = w(k+1) - w(k)$ as before.

Lemma 5.3 *As before, let $m \in \mathbf{N}_+$, $1/3 < \alpha < 1/2$, and $\alpha-(1/4m) > 1/3$. Furthermore, we assume that $4m - 8m\alpha - 1 = 8m(1/2 - \alpha) - 1 > 0$. Then, for any $1 < q < \infty$ and $r \in \mathbf{N}$,*

$$\sum_{k=1}^{\infty} \mathrm{c}_{q,r}(\mathcal{B}(1)^k_{2\alpha,2m}) \leq M_{q,r} \sum_{k=1}^{\infty} k^{2m+\varepsilon} \Big\| \|J[z(k), z(k)]\|^{2m}_{2\alpha,2m-B} \Big\|_{q,r} < \infty, \tag{5.12}$$
$$\sum_{k=1}^{\infty} \mathrm{c}_{q,r}(\mathcal{B}(2)^k_{2\alpha,2m}) \leq M_{q,r} \sum_{k=1}^{\infty} k^{2m+\varepsilon} \Big\| \|J[z(k), w(k)]\|^{2m}_{2\alpha,2m-B} \Big\|_{q,r} < \infty, \tag{5.13}$$
$$\sum_{k=1}^{\infty} \mathrm{c}_{q,r}(\mathcal{B}(3)^k_{2\alpha,2m}) \leq M_{q,r} \sum_{k=1}^{\infty} k^{2m+\varepsilon} \Big\| \|W(k)^1 \otimes Z(k)^1\|^{2m}_{2\alpha,2m-B} \Big\|_{q,r} < \infty. \tag{5.14}$$

Here, $M_{q,r} > 0$ is the constant given in Lemma 4.4. As a result, $\limsup_{k\to\infty} \mathcal{B}(i)^k_{2\alpha,2m}$ is slim for all $i = 1, 2, 3$, which implies quasi-sure convergence of $\{W(k)^2\}_k$.

Proof. Substituting $x = w(k+1)$ and $y = w(k)$ in (5.8), we have

$$W(k+1)^2 - W(k)^2 = J[z(k), z(k)] + J[z(k), w(k)] - J[z(k), w(k)]^* + W(k)^1 \otimes Z(k)^1.$$

If $w \in \cap_{1\leq i\leq 3}(\limsup_{k\to\infty} \mathcal{B}(i)^k_{\alpha,m})^c$, then $\{W(k)^2\}_{k=1}^{\infty}$ is a Cauchy sequence in $(2\alpha, 2m)$-Besov norm. The left inequalities in (5.12)–(5.14) are obvious from Lemma 4.4 of Chebyshev type. By Lemma 4.3 of Borel-Cantelli type, summability in (5.12)–(5.14) implies that $\limsup_{k\to\infty} \mathcal{B}(i)^k_{\alpha,m}$ are slim for all i.

Hence, it is sufficient to show that the sums on the right hand sides of (5.12)–(5.14) converge. First, we consider (5.12). It is not difficult to see that, for a positive constant C_d,

$$\mathbb{E}[|J[z(k), z(k)]_{s,t}|^2] \leq C_d\Big(\frac{1}{2^{2k}} \wedge (t-s)^2\Big), \qquad (s,t) \in \triangle \tag{5.15}$$

A rough sketch of proof of this estimate is as follows. Suppose that $(j-1)/2^k \leq s \leq j/2^k$ and $(l-1)/2^k \leq t \leq l/2^k$ with $j \leq l$. If $j = l$, then we have from (5.3) that

$$J[z(k), z(k)]_{s,t} = \frac{2^{2k}}{2}\Big\{(t - \frac{j-1}{2^k}) \wedge (\frac{j}{2^k} - t) - (s - \frac{j-1}{2^k}) \wedge (\frac{j}{2^k} - s)\Big\}^2$$
$$\times \Big(2w(\frac{2j-1}{2^{k+1}}) - w(\frac{j-1}{2^k}) - w(\frac{j}{2^k})\Big)^{\otimes 2}.$$

In this case, (5.15) is easy. If $j < l$, we use Chen's identity to obtain

$$J[z(k), z(k)]_{s,t} = J[z(k), z(k)]_{s,j/2^k} + J[z(k), z(k)]_{(l-1)/2^k,t} + Z(k)^1_{s,j/2^k} \otimes Z(k)^1_{(l-1)/2^k,t}.$$

In the same way as above, we can estimate the right hand side. Thus, we have shown (5.15).

With (5.15) in hand, we can calculate in the same way as in (5.5) to obtain

$$\Big\| \|J[z(k), z(k)]\|^{4m}_{2\alpha,2m-B} \Big\|_2 \le C_{\alpha,m,d} \Big(\frac{1}{2^{(4m-8m\alpha-1)/2}} \Big)^k.$$

From this and Proposition 4.1, we can easily see that, if $4m - 8m\alpha - 1 > 0$, the sum in (5.12) converges.

Next, we consider (5.14). In this case we have

$$\mathbb{E}[|W(k)^1_{s,t} \otimes Z(k)^1_{s,t}|^2] \le C_d \big(\frac{1}{2^k} \wedge (t-s)^2 \big), \qquad (s,t) \in \triangle \tag{5.16}$$

for some positive constant C_d. From (5.4) and an estimate that $\mathbb{E}[|W(k)^1_{s,t}|^2] \le d(t-s)$ for all k, we can easily show (5.16).

From this and Proposition 4.1 for $W(k)^1_{s,t} \otimes Z(k)^1_{s,t} \in \hat{\mathcal{C}}_2$, we have

$$\begin{aligned}
\Big\| \|W(k)^1 \otimes Z(k)\|^{2m}_{2\alpha,2m-B} \Big\|_2^2 &\le \mathbb{E}\Big[\Big(\iint_{0\le s\le t\le 1} \frac{|W(k)^1_{s,t} \otimes Z(k)^1_{s,t}|^{2m}}{(t-s)^{1+4m\alpha}} dsdt \Big)^2 \Big] \\
&\le \frac{1}{2} \iint_{0\le s\le t\le 1} \frac{\mathbb{E}[|W(k)^1_{s,t} \otimes Z(k)^1_{s,t}|^{4m}]}{(t-s)^{2+8m\alpha}} dsdt \\
&\le C_{m,d} \iint_{0\le s\le t\le 1} \frac{(2^{-k})^{2m} \wedge (t-s)^{4m}}{(t-s)^{2+8m\alpha}} dsdt \\
&\le C_{\alpha,m,d} \Big(\frac{1}{2^{(4m-8m\alpha-1)/2}} \Big)^k.
\end{aligned} \tag{5.17}$$

Here, we used (5.6) for the last inequality.

Applying Proposition 4.1 to $\|W(k)^1 \otimes Z(k)\|^{2m}_{2\alpha,2m-B} \in \hat{\mathcal{C}}_{4m}$, we obtain from (5.17) that

$$\Big\| \|W(k)^1 \otimes Z(k)\|^{2m}_{2\alpha,2m-B} \Big\|_{q,r} \le C_{q,r,\alpha,m,d} \Big(\frac{1}{2^{(4m-8m\alpha-1)/4}} \Big)^k$$

for some positive constant $C_{q,r,\alpha,m,d}$ independent of k. Hence, if $4m - 8m\alpha - 1 > 0$, the right hand side of (5.14) is summable.

Finally, we consider (5.13). In this case we have

$$\mathbb{E}[|J[z(k), w(k)]_{s,t}|^2] \le C_d \big(\frac{1}{2^k} \wedge (t-s)^2 \big), \qquad (s,t) \in \triangle \tag{5.18}$$

for some positive constant C_d. The right hand side of (5.18) is the same as in (5.16). Hence, (5.18) implies summability of (5.13).

Now we give a sketch of proof of (5.18). If $(j-1)/2^k \le s \le t \le j/2^k$ for some j, then we can easily verify (5.18) by using (5.4). If $(j-1)/2^k \le s \le j/2^k$ and $(l-1)/2^k \le t \le l/2^k$ with $j < l$, then we can verify (5.18) by using Chen's identity as follows;

$$\begin{aligned} J[z(k), w(k)]_{s,t} = J[z(k), w(k)]_{s,j/2^k} + \sum_{j+1 \le i \le l-1} J[z(k), w(k)]_{(i-1)/2^k, i/2^k} \\ + J[z(k), w(k)]_{(l-1)/2^k, t} + Z(k)^1_{s,j/2^k} \otimes W(k)^1_{j/2^k, t}. \end{aligned}$$

From this we can see (5.18). Thus, we have shown the lemma. ∎

As a corollary, we have quasi-continuity of Brownian rough path, which was first given by Aida [1] in a slightly stronger form.

Corollary 5.4 *Let α and m be as in Lemmas 5.2 and 5.3. Then, the mapping $w \mapsto W = S_2(w)$ is ∞-quasi continuous.*

Proof. We will show $S_2 : \mathcal{W} \to G\Omega^B_{\alpha,4m}(\mathbf{R}^d)$ is quasi-continuous. Let $n \in \mathbf{N}_+$ be arbitrary. Then, we can find $N = N(n)$ such that

$$\sum_{k=N}^{\infty} \mathrm{c}_{n,n}(\mathcal{A}^k_{\alpha,4m}) < \frac{1}{4n}, \quad \max_{1 \le i \le 3} \sum_{k=N}^{\infty} \mathrm{c}_{n,n}(\mathcal{B}(i)^k_{2\alpha,2m}) < \frac{1}{4n}.$$

On $(\cup_{k=N}^{\infty} \mathcal{A}^k_{\alpha,4m})^c = \cap_{k=N}^{\infty} (\mathcal{A}^k_{\alpha,4m})^c$, $\{W^1(k)\}_k$ converges uniformly as $k \to \infty$. For each k, $w \mapsto W^1(k)$ is clearly continuous. Hence, outside $\cup_{k=N}^{\infty} \mathcal{A}^k_{\alpha,4m}$, $W^1 = S_2(w)^1$ is continuous in w. Similarly, outside $\cup_{i=1}^{3} \cup_{k=N}^{\infty} \mathcal{B}(i)^k_{\alpha,4m}$, $W^2 = S_2(w)^2$ is continuous in w.

Therefore, the set $\cup_{k=N}^{\infty} \mathcal{A}^k_{\alpha,4m} \cup [\cup_{i=1}^{3} \cup_{k=N}^{\infty} \mathcal{B}(i)^k_{\alpha,4m}]$ has (n,n)-capacity smaller than $1/n$, due to subadditivity of the capacity, and its complement is a subset of $\mathcal{Z}_{\alpha,4m}$. Moreover, on its complement, S_2 is continuous. This proves the corollary. ∎

Remark 5.5 *If we assume*

$$m \in \mathbf{N}_+, \quad \frac{1}{3} < \alpha < \frac{1}{2}, \quad \alpha - \frac{1}{4m} > \frac{1}{3}, \quad \textit{and} \quad 4m - 8m\alpha > 2, \tag{5.19}$$

then all the assumptions for α, m in all the results in this section are satisfied. The condition (5.19) might not be best possible, but is sufficient for our purpose. Because what is really needed is as follows; for any $\alpha \in (1/3, 1/2)$, sufficiently large $m \in \mathbf{N}_+$ satisfies the assumptions of these results.

6 Quasi-sure analysis for solutions of SDE

We study SDE (2.1), which involves a small positive parameter $\varepsilon \in (0,1]$. In this section ε is fixed, however. Hence, readers who like it simple may set $\varepsilon = 1$.

Next we introduce a rough differential equation (RDE) associated with the vector fields V_i's and express y^ε in terms of the Lyons-Itô map. For a while we assume V_i's are only C_b^3. Set $\lambda \in C_0^{1-H}([0,1], \mathbf{R})$ by $\lambda_t = t$. For $1/3 < \alpha < 1/2$, we denote by $\tau_\lambda : G\Omega_\alpha^H(\mathbf{R}^d) \to G\Omega_\alpha^H(\mathbf{R}^{d+1})$ the Young pairing formally given by $X \mapsto (X, \lambda)$. We denote by $\Phi_\varepsilon : G\Omega_\alpha^H(\mathbf{R}^{d+1}) \to G\Omega_\alpha^H(\mathbf{R}^n)$ the Lyons-Itô map associated with the coefficients $[V_1, \ldots, V_d, V_0(\varepsilon, \cdot\,)]$, which is continuous by Lyons's continuity theorem. If $X = S_2(x)$ for a 1-Hölder continuous path, then $a + \Phi_\varepsilon(\tau_\lambda(X))^1$ solves the following ODE in Riemann-Stieltjes sense;

$$dz_t = \sum_{i=1}^d V_i(z_t)dx_t^i + V_0(\varepsilon, z_t)dt, \qquad z_0 = a.$$

Hence, by Wong-Zakai's approximation theorem, $y^\varepsilon = a + \Phi_\varepsilon(\tau_\lambda(\varepsilon W))^1$ for a.a.w. Thus, the solution of the scaled SDE is obtained via rough path theory. Note also that this shows y^ε is a $C^{\alpha-H}([0,1], \mathbf{R}^n)$-valued Wiener functional.

Proposition 6.1 *Assume that $[V_1, \ldots, V_d, V_0(\varepsilon, \cdot\,)]$ are C_b^3 and let $(y_t^\varepsilon)_{0\le t\le 1}$ be the solution of SDE (2.1). Then, for any $1/3 < \alpha < 1/2$, the mapping*

$$\mathcal{W} \ni w \mapsto y^\varepsilon \in C^{\alpha-H}([0,1], \mathbf{R}^n)$$

admits an ∞-quasi-redefinition, which is explicitly given by $a + \Phi_\varepsilon(\tau_\lambda(\varepsilon W))^1$.

Proof. For given α, we can find (α', m) as in (5.19) such that $\alpha < \alpha'$ and $\alpha' - (4m)^{-1} \ge \alpha$. By Proposition 3.1, $G\Omega^B_{\alpha',4m}(\mathbf{R}^d)$ is continuously embedded in $G\Omega_\alpha^H(\mathbf{R}^d)$. To keep notations simple, we will not write this injection explicitly.

For all $w \in \mathcal{W}$ and $k \in \mathbf{N}_+$, we have $y(k)^\varepsilon = a + \Phi_\varepsilon(\tau_\lambda(\varepsilon W(k)))^1$, where $y(k)^\varepsilon$ is the unique solution of

$$dy(k)_t^\varepsilon = \sum_{i=1}^d V_i(y(k)_t^\varepsilon)\varepsilon dw(k)_t + V_0(\varepsilon, y(k)_t^\varepsilon)dt, \qquad y(k)_0^\varepsilon = a.$$

Since we assume C_b^3, we may use Wong-Zakai's theorem. It says $\sup_{0\le t\le 1} |y(k)_t^\varepsilon - y_t^\varepsilon| \to 0$ almost surely as $k \to \infty$. On the other hand, by Lyons's continuity theorem, Proposition 3.1, and Corollary 5.4, $\Phi_\varepsilon(\tau_\lambda(\varepsilon W(k)))^1$ converges to $\Phi_\varepsilon(\tau_\lambda(\varepsilon W))^1$ in α-Hölder norm not just almost surely, but also quasi-surely. Hence, we have $y^\varepsilon = a + \Phi_\varepsilon(\tau_\lambda(\varepsilon W))^1$ a.s. Again by Lyons's continuity theorem, Proposition 3.1, and Corollary 5.4, the right hand side is ∞-quasi-continuous in w. This implies that $a + \Phi_\varepsilon(\tau_\lambda(\varepsilon W))^1$ is an ∞-quasi-redefinition of y^ε. ∎

Remark 6.2 *J. Ren [20] studied the quasi-sure refinement of Wong-Zakai's approximation theorem. The proof of the above theorem is a new proof of Ren's result. In fact, it is an improvement since we only assumed C_b^3, not C_b^∞ for the coefficients.*

Remark 6.3 *The above proposition can be regarded as a new proof of Malliavin-Nualart's result in [18]. Here, we give a quick remark about it. Assume that the coefficients V_i's are C_b^∞. Then, for each t, $y_t^\varepsilon \in \mathbf{D}_\infty(\mathbf{R}^n)$ and therefore it admits an ∞-quasi-redefinition by a well-known general theory. J. Ren [20] proved that there exists a nice modification $\hat{y}^\varepsilon$ of y^ε such that, for each t, $\hat{y}_t^\varepsilon$ is an ∞-quasi-redefinition of y_t^ε. (This process $\hat{y}^\varepsilon$ is the limit of Ren's refinement of Wong-Zakai's theorem in [20].)*

But, it is much more difficult to prove the path space-valued functional y^ε admits an ∞-quasi-redefinition. Malliavin and Nualart [18] proved that $w \mapsto y^\varepsilon$ admits an ∞-quasi-redefinition. (More precisely, they proved that Wiener functional

$$\mathcal{W} \ni w \mapsto \Big[[0,1] \times \mathbf{R}^n \ni (t,a) \mapsto y_t^\varepsilon = y^\varepsilon(t,a)\Big] \tag{6.1}$$

which takes values in a space of two-parameter functions, admits an ∞-quasi-redefinition. But, this is not the point here.)

Their method is as follows. Firstly, they introduce a UMD Banach space with certain nice properties and showed y^ε takes values in it. Secondly, they extended Malliavin calculus so that it applies to Wiener functionals which takes values in the UMD Banach space. Thirdly, they regarded SDE (2.1) as an equation that takes values in the UMD Banach space, took $\mathcal{H}$-derivative of it to prove that y^ε is a UMD Banach space-valued $\mathbf{D}_\infty$-Wiener functional. Then, the same argument as in the case of real-valued $\mathbf{D}_\infty$-Wiener functionals applies to y^ε, too, and existence of ∞-redefinition is shown.

Note that, in their method, smoothness of the coefficients needs to be assumed. In our Proposition 6.1, however, we only assumed C_b^3, which is probably astonishing if we do not know rough path theory. In this sense, this is an improvement of Malliavin-Nualart's result.

Since Lyons-Itô map is continuous in the initial value a, too, it is clearly possible to extend Proposition 6.1 to the case of (6.1), by introducing certain Hölder-type norm on the two-parameter function space. It is not clear, however, that we have really improved the results in [18] or not, mainly because we do not know strong the UMD Banach norm is, compared to the Hölder-type norm.

Now, we discuss pinned diffusion measures. By using the arguments we developed so far, we can obtain pinned diffusion processes via rough path theory. This result was first proved in Inahama [12], but the argument here is much more straight forward, due to quasi-continuity of W (Corollary 5.4).

From now on we assume V_i $(0 \le i \le d)$ are C_b^∞. We assume further that the vector fields $\{V_1, \ldots, V_d; V_0^\varepsilon\}$ satisfy hypoellipticity condition at any a. Here, $V_0^\varepsilon = V_0(\varepsilon, \cdot)$. Then, the solution $y_t^\varepsilon = y^\varepsilon(t,a)$ of SDE (2.1) has a density $p_t^\varepsilon(a,a')$ for all $t > 0$, $\varepsilon > 0$,

and $a \in \mathbf{R}^n$, that is, for any Borel set $A \subset \mathbf{R}^n$, $\mathbb{P}(y^\varepsilon(t,a) \in A) = \int_A p_t^\varepsilon(a,a')da'$. By Watanabe's theory, $p_t^\varepsilon(a,a') = \mathbb{E}[\delta_{a'}(y^\varepsilon(t,a))]$. (See Section 5.9 –5.10, [11])

Take any a, a' such that $p_1^\varepsilon(a,a') > 0$. Let $\mathbb{Q}_{a,a'}^\varepsilon$ be the pinned diffusion measure from a to a' associated with the second order differential operator $\mathcal{L}^\varepsilon = (\varepsilon^2/2)\sum_{i=1}^d V_i^2 + V_0^\varepsilon$. This is a probability measure on $C_{a,a'}([0,1],\mathbf{R}^n) := \{x \in C([0,1],\mathbf{R}^n) \mid x_0 = a, x_1 = a'\}$ characterized by the following equation; for any $l \in \mathbf{N}_+$, any partition $\{0 = t_0 < t_1 < \cdots < t_l < t_{l+1} = 1\}$ of $[0,1]$, and any $f \in \mathcal{S}(\mathbf{R}^{nl})$,

$$\int f(x_{t_1},\ldots,x_{t_l})\mathbb{Q}_{a,a'}^\varepsilon(dx) = p_1^\varepsilon(a,a')^{-1}\int_{(\mathbf{R}^n)^l} f(a_1,\ldots,a_l)\prod_{j=1}^{l+1} p_{t_j - t_{j-1}}^\varepsilon(a_{j-1},a_j)\prod_{j=1}^{l} da_j.$$

Here, $a_0 = a$ and $a_{l+1} = a'$ by convention. Actually, by Proposition 6.4 below, $\mathbb{Q}_{a,a'}^\varepsilon$ turns out to be a probability measure on $C_{a,a'}^{\alpha-H}([0,1],\mathbf{R}^n)$ for any $1/3 < \alpha < 1/2$.

Let $\theta_{a,a'}^\varepsilon$ be a Borel finite measure on $\mathcal{W}$ that corresponds to $\delta_{a'}(y^\varepsilon(1,a))$ via Sugita's theorem (Proposition 4.5). We denote by $\hat{\theta}_{a,a'}^\varepsilon$ its normalization. Let α, m be as in (5.19). We set $\mu_{a,a'}^\varepsilon = (\varepsilon S_2)_*\theta_{a,a'}^\varepsilon$ and $\hat{\mu}_{a,a'}^\varepsilon = (\varepsilon S_2)_*\hat{\theta}_{a,a'}^\varepsilon$. Since $\mathcal{Z}_{\alpha,4m}^c$ is slim, the lift map εS_2 is well-defined with respect to these measures. Thus, we obtained measures on $G\Omega_{\alpha,4m}^B(\mathbf{R}^n)$. It is easy to see that the support of $\hat{\mu}_{a,a'}^\varepsilon$ is actually contained in the closed subset $\{X \mid a + \Phi_\varepsilon(\tau_\lambda(X)) = a'\}$.

Proposition 6.4 *Keep all the notations and assumptions as above. Then, the push-forward measure $[(a+\Phi_\varepsilon^1)\circ\tau_\lambda]_*\hat{\mu}_{a,a'}^\varepsilon$ is the pinned diffusion measure $\mathbb{Q}_{a,a'}^\varepsilon$. Here, Φ_ε^1 denotes the first level of the Itô map Φ_ε.*

Proof. For simplicity, we write $\tilde{y}^\varepsilon = a + \Phi_\varepsilon(\tau_\lambda(\varepsilon S_2(w)))^1$ as in the proof of Proposition 6.1. By Proposition 6.1, this is an ∞-quasi-redefinition of y^ε. It is sufficient to prove that $[(a+\Phi_\varepsilon^1)\circ\tau_\lambda\circ\varepsilon S_2]_*\theta_{a,a'}^\varepsilon = p_1^\varepsilon(a,a')\cdot\mathbb{Q}_{a,a'}^\varepsilon$. This fact is well-known is quasi-sure analysis, but we give a proof here for readers' convenience.

Let $\chi : \mathbf{R}^n \to \mathbf{R}$ be C^∞ with compact support such that $\chi \geq 0$ and $\int_{\mathbf{R}^n}\chi(\xi)d\xi = 1$. For $k \in \mathbf{N}_+$, set $\psi_k(\xi) = k^n\chi(k(\xi - a'))$. Then, for any function $g : \mathbf{R}^n \to \mathbf{R}$ which is continuous around a', we have $\lim_{k\to\infty}\int_{\mathbf{R}^n}\psi_k(\xi)g(\xi)d\xi = g(a')$. In particular, $\lim_{k\to\infty}\psi_k = \delta_{a'}$ in $\mathcal{S}'(\mathbf{R}^n)$.

Take any partition $\{0 = t_0 < t_1 < \cdots < t_l < t_{l+1} = 1\}$ of $[0,1]$ and any $f \in \mathcal{S}(\mathbf{R}^{nl})$.

By Proposition 4.5, we have

$$
\begin{aligned}
\int_{\mathcal{W}} f(\tilde{y}^{\varepsilon}_{t_1},\ldots,\tilde{y}^{\varepsilon}_{t_l})\theta^{\varepsilon}_{a,a'}(dw) &= \mathbb{E}\big[f(y^{\varepsilon}_{t_1},\ldots,y^{\varepsilon}_{t_l})\cdot\delta_{a'}(y^{\varepsilon}_{t})\big] \\
&= \lim_{k\to\infty}\mathbb{E}\big[f(y^{\varepsilon}_{t_1},\ldots,y^{\varepsilon}_{t_l})\psi_k(y^{\varepsilon}_{t})\big] \\
&= \lim_{k\to\infty}\int_{(\mathbf{R}^n)^{l+1}} f(a_1,\ldots,a_l)\psi_k(a_{l+1})\prod_{j=1}^{l+1}p^{\varepsilon}_{t_j-t_{j-1}}(a_{j-1},a_j)\prod_{j=1}^{l+1}da_j. \\
&= \int_{(\mathbf{R}^n)^{l}} f(a_1,\ldots,a_l)\prod_{j=1}^{l}p^{\varepsilon}_{t_j-t_{j-1}}(a_{j-1},a_j)p^{\varepsilon}_{t_{l+1}-t_l}(a_l,a')\prod_{j=1}^{l}da_j.
\end{aligned}
$$

where $a_0 = a$. This completes the proof. ▮

7 Large deviation for $\mu^{\varepsilon}_{a,a'}$ on geometric rough path space

The aim of this section is to prove that the family $\{\mu^{\varepsilon}_{a,a'}\}_{\varepsilon>0}$ of finite measures on $G\Omega^{B}_{\alpha,4m}$ satisfies a large deviation principle as $\varepsilon\searrow 0$ with a good rate function. We assume the coefficient of SDE (2.1) are C^{∞}_{b}. Note that in the most important theorem, Theorem 7.1 (i), we will assume the ellipticity condition only at the starting point a.

By Lyons's continuity theorem and the contraction principle (e.g., Thorem 4.2.1, [2]), a large deviation principle for the pushforward of $\mu^{\varepsilon}_{a,a'}$ by the map $\tilde{y}^{\varepsilon}$ immediately follows (see Corollary 7.2 below). This can be considered as a special case of Theorem 2.1, [23]. After normalization, our main theorem, i.e., a large deviation principle for $\{\mathbb{Q}^{\varepsilon}_{a,a'}\}_{\varepsilon>0}$ under **(H1)**, is easily obtained.

Let $\phi^0 = \phi^0(h)$ be the unique solution of ODE (2.2). Note that $\phi^0(h) = a + \Phi_0(\tau_\lambda(S_2(h)))^1$. Under **(H1)**, $\mathcal{K}^{a,a'} = \{h\in\mathcal{H} \mid \phi^0(h)_1 = a'\}$ is a non-empty closed subset of $\mathcal{H}$. In general, however, it can be empty.

For α and m as in (5.19), we set a rate function $I : G\Omega^{B}_{\alpha,4m}(\mathbf{R}^n)\to[0,\infty]$ as follows;

$$
I(X) = \begin{cases} \|h\|^2_{\mathcal{H}}/2 & (\text{if } X = S_2(h) \text{ for some } h\in\mathcal{K}^{a,a'}), \\ \infty & (\text{otherwise}). \end{cases}
$$

We also set $\hat{I}(X) = I(X) - \min\{\|h\|^2_{\mathcal{H}}/2 \mid h\in\mathcal{K}^{a,a'}\}$. Note here that the minimum is actually attained.

Recall that a Schilder-type large deviation for the laws of scaled Brownian rough path εW holds on $G\Omega^{H}_{\alpha'}(\mathbf{R}^d)$ as $\varepsilon\searrow 0$ for any $1/3<\alpha'<1/2$ with a good rate function I_{Sch} (see Friz-Victoir [6]). Here, $I_{Sch}(X) = \|h\|^2_{\mathcal{H}}/2$ if $X = S_2(h)$ for some $h\in\mathcal{H}$ and $I_{Sch}(X)=\infty$ if otherwise. For $\alpha<\alpha'<1/2$, there is a continuous embedding

$G\Omega^H_{\alpha'}(\mathbf{R}^d) \hookrightarrow G\Omega^B_{\alpha,4m}(\mathbf{R}^d)$. Hence, by the contraction principle, the Schilder-type large deviation also holds on $G\Omega^B_{\alpha,4m}(\mathbf{R}^d)$ and I_{Sch} is still good on $G\Omega^B_{\alpha,4m}(\mathbf{R}^d)$. We can easily verify from this that I defined above is lower semicontinuous and good (i.e. the level set $\{X | I(X) \le c\}$ is compact for all $0 \le c < \infty$). So is $\hat{I}$.

The following theorem is the key result in this paper, which may be more important than our main theorem (Theorem 2.1). The condition on α, m is so that all the results in the previous sections are available. An advantage of the formulation in Theorem 7.1, (i) below is that we do not need to worry about whether the pinned diffusion measures are well-defined or not. (Note that $\mu^\varepsilon_{a,a'}$ can be the zero measure in the worst case.)

Theorem 7.1 *We keep the same notations and assumptions as above. Let $m \in \mathbf{N}_+$ and $1/3 < \alpha < 1/2$ be as in (5.19). Then, the following* (i)–(ii) *hold:*
(i) *Assume that $\{V_1(a), \dots, V_d(a)\}$ linearly spans $\mathbf{R}^n = T_a\mathbf{R}^n$ at the starting point a. The family $\{\mu^\varepsilon_{a,a'}\}_{\varepsilon>0}$ of finite measures on $G\Omega^B_{\alpha,4m}(\mathbf{R}^d)$ satisfies a large deviation principle as $\varepsilon \searrow 0$ with a good rate function I, that is, for any Borel set $A \subset G\Omega^B_{\alpha,4m}(\mathbf{R}^d)$, the following inequalities hold;*

$$- \inf_{X \in A^\circ} I(X) \le \liminf_{\varepsilon \searrow 0} \varepsilon^2 \log \mu^\varepsilon_{a,a'}(A) \le \limsup_{\varepsilon \searrow 0} \varepsilon^2 \log \mu^\varepsilon_{a,a'}(A) \le - \inf_{X \in \bar{A}} I(X).$$

(ii) *Assume* **(H1)** *in addition. Then, the family $\{\hat{\mu}^\varepsilon_{a,a'}\}_{\varepsilon>0}$ of probability measures on $G\Omega^B_{\alpha,4m}(\mathbf{R}^d)$ satisfies a large deviation principle as $\varepsilon \searrow 0$ with a good rate function $\hat{I}$.*

Under **(H1)** , the heat kernel $p^\varepsilon_t(a,a')$ exists and positive for all $a, a' \in \mathbf{R}^n$, $t > 0$ and $\varepsilon > 0$. Hence, the results in the previous section are available. Hence, the second assertion is immediate from the first one, because the weight of the whole set equals $p^\varepsilon_1(a,a')$ and the whole set is both open and closed. The proof of the first assertion of Theorem 7.1 will be given in later sections.

As an immediate corollary of Theorem 7.1, we can show our main thorem (Theorem 2.1), i.e., a large deviation principle for the family $\{\mathbb{Q}^\varepsilon_{a,a'}\}_{\varepsilon>0}$ of pinned diffusion measures. The key point is that we can use the contraction principle, thanks to Lyons's continuity theorem.

Corollary 7.2 *Keep the same notations and assumptions as in Theorem 7.1, (i). Then, the pushforward measure of $\mu^\varepsilon_{a,a'}$ by the map $\tilde{y}^\varepsilon$ satisfies a large deviation principle with a good rate function $J'(y)$, where $J'(y) = \inf\{\|h\|^2_{\mathcal{H}}/2 \mid h \in \mathcal{K}^{a,a'}$ with $y = \phi^0(h)\}$ if $y = \phi^0(h)$ for some $h \in \mathcal{K}^{a,a'}$ and $J'(y) = \infty$ if no such $h \in \mathcal{K}^{a,a'}$ exists.*

Proof of Corollary 7.2 and Theorem 2.1. There exists $\alpha' \in (\alpha, 1/2)$ and $m \in \mathbf{N}_+$ such that $\alpha' - (4m)^{-1} \ge \alpha$ and (5.19) hold. Then, we can use Theorem 7.1 above and we have a continuous embedding $G\Omega^B_{\alpha',4m}(\mathbf{R}^d) \hookrightarrow G\Omega^H_{\alpha}(\mathbf{R}^d)$ by Proposition 3.1. Of course, the Young pairing $\tau_\lambda : G\Omega^H_{\alpha}(\mathbf{R}^d) \to G\Omega^H_{\alpha}(\mathbf{R}^{d+1})$ is continuous.

Let us give a remark on continuity of the Lyons-Itô map Φ_ε since it depends on ε, too. The coefficient V_0 for the drift term depends on ε and $C_b^\infty([0,1]\times\mathbf{R}^n)$. Hence, by Taylor expansion, we have

$$\sup_{\xi:|\xi|\le R}|\nabla^i_{(\xi)}V_0(\varepsilon,\xi)-\nabla^i_{(\xi)}V_0(\varepsilon_0,\xi)|\to 0 \qquad \text{as } \varepsilon\to\varepsilon_0$$

for any $R>0$ and $0\le i\le 3$. Here, $\nabla_{(\xi)}$ stands for the gradient in ξ-variable.

Under this kind uniform convergence of the ε-dependent coefficient, a slight generalization of Lyons's continuity theorem holds, which claims that the mapping

$$[0,1]\times G\Omega^H_\alpha(\mathbf{R}^{d+1})\ni(\varepsilon,Z)\mapsto\Phi_\varepsilon(Z)\in G\Omega^H_\alpha(\mathbf{R}^n)$$

is continuous. It immediately follows that the mapping

$$[0,1]\times G\Omega^B_{\alpha',4m}(\mathbf{R}^d)\ni(\varepsilon,X)\mapsto a+\Phi_\varepsilon(\tau_\lambda(X))^1\in C_a^{\alpha-H}([0,1],\mathbf{R}^d)$$

is continuous, too. Now we can use Proposition 6.4 and a slight generalization of the contraction principle for large deviations (see Lemma 3.9, [13] for example). This completes the proof. ∎

8 Proof of Theorem 7.1 (i): Lower estimate

The aim of this section is to prove the lower estimate in Theorem 7.1 (i).

Before we start our proof, let us rewrite SDE (2.1) in matrix form. Set $\sigma:\mathbf{R}^n\to\mathrm{Mat}(n,d)$ by $[V_1,\ldots,V_d]$, that is, the jth column vector of σ is V_j. We also write $V_0(\varepsilon,\xi)=b(\varepsilon,\xi)$. Then, SDE (2.1) can be rewritten as follows;

$$dy^\varepsilon_t=\sigma(y^\varepsilon_t)\circ\varepsilon dw_t+b(\varepsilon,y^\varepsilon_t)dt,\qquad y^\varepsilon_0=a. \tag{8.1}$$

This matrix form (8.1) is sometimes more convenient than (2.1). In the same way, ODE (2.2) is equivalent to the following ODE;

$$d\phi^0_t=\sigma(\phi^0_t)dh_t+b(0,\phi^0_t)dt,\qquad \phi^0_0=a.$$

Let $(y^\varepsilon_t(w))_{0\le t\le 1}$ be the solution of SDE (8.1). For $h\in\mathcal{H}$, we consider the Cameron-Martin shift of y^ε.

$$dy^{\varepsilon,h}_t=\sigma(y^{\varepsilon,h}_t)[\circ\varepsilon dw_t+dh_t]+b(\varepsilon,y^{\varepsilon,h}_t)dt,\qquad y^{\varepsilon,h}_0=a. \tag{8.2}$$

An ∞-quasi-continuous version of $y^{\varepsilon,h}$ is given by $a+\Phi_\varepsilon(\tau_\lambda(S_2(\varepsilon w+h)))$ by Lemma 5.1 and Proposition 6.1.

Non-degeneracy of the Malliavin covariance matrix of the solution of SDE with under ellipticity condition is in Section 6.1.5, Shigekawa [21] for example. By repeating the same proof with carefully taking ε-dependency into account, we can verify that $(y^{\varepsilon,h}_1-a')/\varepsilon$

is uniformly non-degenerate in the sense of Malliavin as $\varepsilon \searrow 0$. (This fact was already shown in [24] when the drift term is of the form $b(\varepsilon, \xi) = \varepsilon^2 \hat{b}(\xi)$ for an ε-independent vector field $\hat{b}$. The definition of uniform non-degeneracy is in p. 10, [24]. When we try to extend our method to the hypoelliptic case, this part becomes very difficult.)

Moreover, in the same way as in Watanabe [24], we have the following asymptotics for $y_1^{\varepsilon,h}$ as $\varepsilon \searrow 0$;

$$y_1^{\varepsilon,h} = \phi^0(h)_1 + \varepsilon \phi^1(\,\cdot\,; h)_1 + O(\varepsilon^2) \qquad \text{in } \mathbf{D}_\infty(\mathbf{R}^n).$$

Watanabe obtained the expansion up to any order when the drift term is of the form $b(\varepsilon, \xi) = \varepsilon^2 \hat{b}(\xi)$. But we only need an expansion up to second order. Modification for the case of general $b(\varepsilon, \xi)$ is easy. Here, $\phi_t^1 = \phi^1(w; h)_t$ is the unique solution of the following equation;

$$d\phi_t^1 - \nabla\sigma(\phi_t^0)\langle \phi_t^1, dh_t\rangle - \nabla_{(\xi)} b(0, \phi_t^0)\langle \phi_t^1\rangle dt = \sigma(\phi_t^0) dw_t + \partial_\varepsilon b(0, \phi_t^0) dt, \quad \phi_0^1 = 0.$$

Here, $\nabla_{(\xi)}$ denotes the (partial) gradient with respect to the second variable. ∂_ε should be understood in a similar way.

Now we give an explicit expression for ϕ_t^1. Let $M_t = M(h)_t$ be a unique solution of the following $n \times n$ matrix-valued ODE;

$$dM_t = [\nabla\sigma(\phi_t^0)\langle \bullet, dh_t\rangle + \nabla_{(\xi)} b(0, \phi_t^0)\langle \bullet \rangle dt] \cdot M_t, \qquad M_0 = \mathrm{Id}_n.$$

Then, M_t is invertible for all t and we have

$$\phi_t^1 = M_t \int_0^t M_s^{-1}[\sigma(\phi_s^0) dw_s + \partial_\varepsilon b(0, \phi_s^0) ds].$$

Hence, as a functional in w, ϕ_1^1 is an $\mathbf{R}^n$-valued Gaussian random variable with mean $M_1 \int_0^1 M_s^{-1} \nabla_\varepsilon b(0, \phi_s^0) ds$ and covariance $M_1 \int_0^1 M_s^{-1} \sigma(\phi_s^0)\sigma(\phi_s^0)^T (M_s^{-1})^T ds M_1^T$. Since we assumed ellipticity at the initial point a, $\sigma\sigma^T$ is positive and bounded away from 0 near a. Thereore, this covariance matrix is non-degenerate for any h. In particular, $\mathbb{E}[\delta_0(\phi_1^1)] > 0$.

Let $h \in \mathcal{K}^{a,a'}$. By (a special case of) Theorem 2.3, [24] or Theorem 9.4, [11] and the uniform non-degeneracy, we have

$$\lim_{\varepsilon \searrow 0} \delta_0\Big(\frac{y_1^{\varepsilon,h} - a'}{\varepsilon}\Big) = \delta_0(\phi_1^1) \qquad \text{in } \mathbf{D}_{q,-r}(\mathbf{R}^n)$$

for some $q \in (1, \infty)$ and some $r \in \mathbf{N}$.

Lemma 8.1 *$\mathcal{K}^{a,a'} \cap \mathcal{W}^*$ is dense in $\mathcal{K}^{a,a'}$. In other words, for any $h \in \mathcal{K}^{a,a'}$, there exist $h_j \in \mathcal{K}^{a,a'}$ $(j \in \mathbf{N}_+)$ such that $\langle h_j, \cdot\,\rangle_{\mathcal{H}} \in \mathcal{W}^*$ for all j and $\lim_{j\to\infty} \|h_j - h\|_{\mathcal{H}} = 0$.*

Proof. Set $F(h) = \phi^0(h)_1$. Then, $F : \mathcal{H} \to \mathbf{R}^n$ is Fréchet-C^1 and its Fréchet derivative DF is given by

$$D_k F(h) = \langle DF(h), k\rangle_{\mathcal{H}} = M_1 \int_0^1 M_s^{-1}\sigma(\phi^0(h)_s)k'_s ds.$$

Surjectivity of the linear map $DF(h) : \mathcal{H} \to \mathbf{R}^n$ is equivalent to non-degeneracy of the "deterministic Malliavin covariance" matrix

$$(\langle DF^i(h), DF^j(h)\rangle_{\mathcal{H}})_{1\le i,j\le n} = M_1 \int_0^1 M_s^{-1}\sigma(\phi^0(h)_s)\sigma(\phi^0(h)_s)^T(M_s^{-1})^T ds M_1^T,$$

which has already been shown. Hence, we can use Lemma 8.2 below with $\mathcal{K} = \mathcal{H}$ and $\mathcal{L} = \mathcal{W}^*$. ∎

Lemma 8.2 *Let $\mathcal{K}$ be a real Hilbert space and $\xi \in \mathcal{K}$. Assume that F is an $\mathbf{R}^n$-valued Fréchet-C^1 map defined on a neighborhood of ξ with a bounded derivative DF. Let $\mathcal{L}$ be a real Banach space which is continuously and densely embedded in $\mathcal{K}$. Then, there exists $\xi_j \in \mathcal{L}$ $(j = 1, 2, \ldots)$ such that $\lim_{j\to\infty} \|\xi_j - \xi\|_{\mathcal{K}} = 0$ and $F(\xi_j) = F(\xi)$ for all j.*

Proof. Without loss of generality, we may assume $F(\xi) = 0$. It is sufficient to show that, for sufficiently small $r > 0$, there exists $\eta \in \mathcal{L} \cap B_{\mathcal{K}}(\xi, r)$ such that $F(\eta)$=0, where $B_{\mathcal{K}}(\xi, r)$ is the open $\mathcal{K}$-ball of radius r, centered at ξ. There exists an n-dimensional subspace $V \subset \mathcal{K}$ such that $\nabla F(\xi)|_V : V \to \mathbf{R}^n$ is a linear bijection. Hence, we may apply inverse function theorem at ξ to $F|_{\xi+V}$, which is the restriction of F onto the affine subspace $\xi + V$.

First, we consider the case for $n = 1$. For any sufficiently small $r > 0$, there exist $\chi_0, \chi_1 \in (\xi + V) \cap B_{\mathcal{K}}(\xi, r)$ such that $F(\chi_0) > 0$ and $F(\chi_1) < 0$. Since $\mathcal{L}$ is dense in $\mathcal{K}$, we can find $\eta_0, \eta_1 \in \mathcal{L} \cap B_{\mathcal{K}}(\xi, r)$ such that $F(\eta_0) > 0$ and $F(\eta_1) < 0$. Then, the line segment $\overline{\eta_0\eta_1} = \{\tau\eta_0 + (1 - \tau)\eta_1 \mid 0 \le \tau \le 1\}$ is still in $\mathcal{L} \cap B_{\mathcal{K}}(\xi, r)$. Then, by the intermediate value theorem, there exists $\tau' \in (0, 1)$ such that $F(\tau'\eta_0 + (1 - \tau')\eta_1) = 0$. Thus, we have shown the case $n = 1$.

Next, we consider the case for $n = 2$. For $0 < r < r'$, set $B_{\mathbf{R}^n}(0, r) = \{a \in \mathbf{R}^n \mid |a| < r\}$, $S^1_r = \{a \in \mathbf{R}^n \mid |a| = r\}$, and $A^1_{r,r'} = \{a \in \mathbf{R}^n \mid r < |a| < r'\}$. For sufficiently small $r > 0$, $F((\xi + V) \cap B_{\mathcal{K}}(\xi, r))$ is an open neighborhood of 0 and $f := F|_{(\xi+V)\cap B_{\mathcal{K}}(\xi,r)}$ is a diffeomorphism. There exists $\rho > 0$ such that $B_{\mathbf{R}^n}(0, 3\rho) \subset F((\xi + V) \cap B_{\mathcal{K}}(\xi, r))$. Then, there exist $N \in \mathbf{N}_+$ and $\xi_0, \ldots, \xi_N \in S^1_\rho$ such that the following conditions are satisfied;

1. ξ_i and ξ_{i+1} are adjacent in counter-clockwise order (with $\xi_{N+1} = \xi_0$),
2. the arc between ξ_i and ξ_{i+1} can be continuously deformed inside $A^1_{\rho/2,2\rho}$, with the end points being fixed, so that the image of the arc by f^{-1} is deformed to the line segment $\overline{f^{-1}(\xi_i)f^{-1}(\xi_{i+1})}$.

Then, there exists $\eta_0, \ldots, \eta_N \in \mathcal{L} \cap B_{\mathcal{K}}(\xi, r)$ such that the following conditions are satisfied;

1. there exists a continuous deformation of $\cup_{i=0}^{N} \overline{f^{-1}(\xi_i) f^{-1}(\xi_{i+1})}$ inside $B_{\mathcal{K}}(\xi, r)$, which deforms each $\overline{f^{-1}(\xi_i) f^{-1}(\xi_{i+1})}$ to $\overline{\eta_i \eta_{i+1}}$.

2. the image of the above continuous deformation by F stays inside $A^1_{\rho/3,3\rho}$.

Obviously, $\cup_{i=0}^{N} \overline{\eta_i \eta_{i+1}}$ is a subset of the following (possibly degenerate) simplex

$$\triangle := \{\eta_0 + \sum_{i=1}^{N} \tau_i (\eta_i - \eta_0) \mid \tau_i \geq 0 \text{ for all } i \text{ and } \textstyle\sum_{i=1}^{N} \tau_i \leq 1\}.$$

This is the convex hull of $\eta_0, \ldots, \eta_N$. Hence, $\triangle \subset B_{\mathcal{K}}(\xi, r) \cap \mathcal{L}$. It is easy to see that $\triangle$ admits a continuous deformation to a single point set $\{\eta_0\}$ without getting out of $\triangle$.

Combining them all, we have shown that S^1_ρ is continuously deformed in $\mathbf{R}^n$ to a single point set $\{F(\eta_0)\}$. More precisely, there exists a continuous map $\phi : [0,1] \times S^1_\rho \to \mathbf{R}^n$ such that $\phi(0, \cdot) = \mathrm{Id}_{S^1_\rho}$ and $\phi(1, \cdot)$ is constant. By Lemma 8.3 below, there exists $(s, a) \in [0,1] \times S^1_\rho$ such that $\phi(s, a) = 0$. By way of construction of the deformation map, there must be $\eta \in \triangle$ such that $F(\eta) = \phi(s, a) = 0$. This proves the case $n = 2$. The higher dimensional cases can be shown in the same way. ▌

Lemma 8.3 *Let $B^n = \{a \in \mathbf{R}^n \mid |a| \leq 1\}$ be the n-dimensional closed ball and let $S^{n-1} = \{a \in \mathbf{R}^n \mid |a| = 1\}$ be the $(n-1)$-dimensional sphere ($n \geq 2$).*
(i) *Let $f : [0,1] \times S^{n-1} \to \mathbf{R}^n$ be a continuous map such that $f(0, \cdot) = \mathrm{Id}_{S^{n-1}}$ and $f(1, \cdot)$ is a constant map. Then, there exists $(s, a) \in [0,1] \times S^{k-1}$ such that $f(s, a) = 0$.*
(ii) *Let $g : B^n \to \mathbf{R}^n$ be a continuous map such that $g|_{S^{n-1}} = \mathrm{Id}_{S^{n-1}}$. Then, there exists $a \in B^n$ such that $g(a) = 0$.*

Proof. Since $f(1, \cdot)$ is constant, there exists a continuous map $g : B^n \to \mathbf{R}^n$ such that $g|_{S^{n-1}} = \mathrm{Id}_{S^{n-1}}$ and $\mathrm{Im} f = \mathrm{Im} g$. Hence, it is sufficient to prove the second assertion. Assume that $g^{-1}(0) = \emptyset$. Then g is continuous from B^n to $\mathbf{R}^n \setminus \{0\}$ and we have the following commutative diagram:

$$\begin{array}{ccc}
\{0\} = H_{n-1}(B^n) & \xrightarrow{g_*} & H_{n-1}(\mathbf{R}^n \setminus \{0\}) = \mathbf{Z} \\
{\scriptstyle \iota_*}\uparrow & & \uparrow{\scriptstyle \iota_*} \\
\mathbf{Z} = H_{n-1}(S^{n-1}) & \xrightarrow{(g|_{S^{n-1}})_*} & H_{n-1}(S^{n-1}) = \mathbf{Z}
\end{array}$$

where ι stands for the injections. $g_* \circ (\iota : S^{n-1} \hookrightarrow B^n)_*$ is the zero map. On the other hand, $(\iota : S^{n-1} \hookrightarrow \mathbf{R}^n \setminus \{0\})_* \circ (g|_{S^{n-1}})_* = (\iota : S^{n-1} \hookrightarrow \mathbf{R}^n \setminus \{0\})_* \circ \mathrm{id}_{H_{n-1}(S^{n-1})}$ is an isomorphism, due to homoyopy equivalence. This is a contradiction. Therefore, $g^{-1}(0) \neq \emptyset$. ▌

The lower estimate in Theorem 7.1 (i) is almost immediate from Lemma 8.1 above and the following proposition.

Proposition 8.4 *Let $A \subset G\Omega^B_{\alpha,4m}(\mathbf{R}^d)$ be open and suppose that $S_2(h) \in A$ and $\langle h, \cdot\rangle_{\mathcal{H}} \in \mathcal{W}^* \cap \mathcal{K}^{a,a'}$. Then, we have $-\|h\|^2_{\mathcal{H}}/2 \le \liminf_{\varepsilon\searrow 0} \varepsilon^2 \log \mu^{\varepsilon}_{a,a'}(A)$.*

Proof. For $R > 0$, we set $B_R = \{X \in G\Omega^B_{\alpha,4m}(\mathbf{R}^d) \mid \|X^i\|_{i\alpha,4m/i-B} < R^i \quad (i = 1, 2)\}$ and $\hat{B}_R(H) = T_h(B_R)$, where T_h is the Young translation by h on $G\Omega^B_{\alpha,4m}(\mathbf{R}^d)$. Then, $\{\hat{B}_R(H) \mid R > 0\}$ forms a fundamental system of open neighborhood around $H = S_2(h)$. Since A is open, there exists $R_0 > 0$ such that $\hat{B}_{R_0}(H) \subset A$. We will estimate the weight of $\hat{B}_{R_0}(H)$ instead of that of A.

By Cameron-Martin formula, it holds that, for any $F \in \mathbf{D}_\infty$,

$$\begin{aligned}\mathbb{E}[F\delta_{a'}(y^{\varepsilon}_1)] &= \mathbb{E}[\exp(-\frac{\langle h, w\rangle}{\varepsilon} - \frac{\|h\|^2_{\mathcal{H}}}{2\varepsilon^2})F(w + \frac{h}{\varepsilon})\delta_{a'}(y^{\varepsilon,h}_1)] \\ &= e^{-\|h\|^2_{\mathcal{H}}/2\varepsilon^2}\varepsilon^{-n}\mathbb{E}[e^{-\langle h,w\rangle/\varepsilon}F(w + \frac{h}{\varepsilon})\delta_0(\frac{y^{\varepsilon,h}_1 - a'}{\varepsilon})].\end{aligned}$$

Here, we used the fact that $\delta_0(\varepsilon\xi) = \varepsilon^{-n}\delta_0(\xi)$.

We denote by $\nu^\varepsilon = \nu^\varepsilon_{a,a'}$ the Borel measure corresponding to $\delta_0((y^{\varepsilon,h}_1 - a')/\varepsilon)$ via Sugita's theorem (Proposition 4.5). Then, we have

$$\begin{aligned}\mu^{\varepsilon}_{a,a'}(\hat{B}_R(H)) &= \int_{\mathcal{W}} I_{\hat{B}_R(H)}(\varepsilon W)\, \theta^{\varepsilon}_{a,a'}(dw) \\ &= e^{-\|h\|^2_{\mathcal{H}}/2\varepsilon^2}\varepsilon^{-n}\int_{\mathcal{W}} e^{-\langle h,w\rangle/\varepsilon} I_{B_{R/\varepsilon}}(W)\nu^\varepsilon(dw).\end{aligned}$$

for any $0 < R < R_0$. Noting that $|\langle h, w\rangle/\varepsilon| \le \|h\|_{\mathcal{W}^*}\|w/\varepsilon\|_{\mathcal{W}} \le C\|h\|_{\mathcal{W}^*}R/\varepsilon^2$, where $C := \sup_{w\neq 0}\|w\|_\infty/\|w\|_{\alpha,4m-B}$, we see that

$$\mu^{\varepsilon}_{a,a'}(\hat{B}_R(H)) \ge e^{-\|h\|^2_{\mathcal{H}}/(2\varepsilon^2)}\varepsilon^{-n}e^{-C\|h\|_{\mathcal{W}^*}R/\varepsilon^2}\nu^\varepsilon\big(\{\|W^i\|^{1/i}_{i\alpha,4m/i-B} < R/\varepsilon \ (i = 1, 2)\}\big).$$

Hence, it is sufficient to show that, for each fixed R,

$$\lim_{\varepsilon\searrow 0}\nu^\varepsilon\big(\{\|W^i\|^{1/i}_{i\alpha,4m/i-B} < R/\varepsilon \ (i = 1, 2)\}\big) = \nu^\infty(\mathcal{W}) = \mathbb{E}[\delta_0(\phi^1_1)] > 0, \qquad (8.3)$$

where ν^∞ stands for the Borel measure corresponding to $\delta_0(\phi^1_1)$ via Sugita's theorem. Indeed, (8.3) implies that

$$\liminf_{\varepsilon\searrow 0}\varepsilon^2\log\mu^{\varepsilon}_{a,a'}(\hat{B}_{R_0}(H)) \ge \liminf_{\varepsilon\searrow 0}\varepsilon^2\log\mu^{\varepsilon}_{a,a'}(\hat{B}_R(H)) \ge -\frac{\|h\|^2_{\mathcal{H}}}{2} - C\|h\|_{\mathcal{W}^*}R$$

Letting $R \searrow 0$, we have $\liminf_{\varepsilon\searrow 0}\varepsilon^2\log\mu^{\varepsilon}_{a,a'}(\hat{B}_{R_0}(H)) \ge -\|h\|^2_{\mathcal{H}}/2$.

It remains to prove (8.3). Since $\delta_0((y_1^{\varepsilon,h}-a')/\varepsilon) \to \delta_0(\phi_1^1)$ in $\mathbf{D}_{q',-r}$ for some $q' \in (1,\infty)$ and $r \geq 0$, we have $\nu^\varepsilon(\mathcal{W}) - \nu^\infty(\mathcal{W}) = \mathbb{E}[1 \cdot \{\delta_0((y_1^{\varepsilon,h} - a')/\varepsilon) - \delta_0(\phi_1^1)\}] \to 0$ as $\varepsilon \searrow 0$. Now, it is sufficient to show that

$$\nu^\varepsilon\bigl(\{\|W^i\|_{i\alpha,4m/i-B}^{1/i} < R/\varepsilon \ (i=1,2)\}^c\bigr) \leq \sum_{i=1,2} \nu^\varepsilon\bigl(\{\|W^i\|_{i\alpha,4m/i-B}^{1/i} \geq R/\varepsilon\}\bigr)$$
$$\leq \sum_{i=1,2} \|\delta_0((y_1^{\varepsilon,h} - a')/\varepsilon)\|_{q',-r} \cdot \mathrm{c}_{q,r}\bigl(\{\|W^i\|_{i\alpha,4m/i-B}^{1/i} \geq R/\varepsilon\}\bigr) \to 0$$

as $\varepsilon \searrow 0$. Here, $1/q + 1/q' = 1$. Since $\|\delta_0((y_1^{\varepsilon,h} - a')/\varepsilon)\|_{q',-r}$ is bounded in ε, the problem is reduced to showing

$$\lim_{\varepsilon \searrow 0} \mathrm{c}_{q,r}\bigl(\{\|W^i\|_{i\alpha,4m/i-B}^{1/i} \geq R/\varepsilon\}\bigr) = 0 \qquad (i = 1,2).$$

This will be shown in the next lemma below. ∎

On an abstract Wiener space, weight of the complement set of the large ball admits Gaussian decay. This is called a large deviation estimate. A similar fact holds in rough path settings, too (for instance, see [5]). In this paper, however, we need a large deviation estimate for capacities, not for measures.

Lemma 8.5 *Let α and m be as in Theorem 7.1. For any $q \in (1,\infty)$ and $r \in \mathbf{N}$, there exist $c > 0$ and $R_1 > 0$ such that*

$$\mathrm{c}_{q,r}\bigl(\{w \in \mathcal{W} \mid \|W^i\|_{i\alpha,4m/i-B}^{1/i} \geq R\}\bigr) \leq e^{-cR^2} \qquad \text{for any } R \geq R_1 \text{ and } i = 1,2.$$

The constants c and R_1 may depend on α, m, q, r, but not on R.

Proof. First, let us check that there exists $c_1 > 0$ such that

$$\mathbb{P}\bigl(\{w \in \mathcal{W} \mid \|W^i\|_{i\alpha,4m/i-B}^{1/i} \geq R\}\bigr) \leq e^{-c_1 R^2} \tag{8.4}$$

holds for sufficiently large $R > 0$. On the Hölder geometric rough path space $G\Omega_\alpha^H(\mathbf{R}^d)$, with any $1/3 < \alpha < 1/2$, this type of Gaussian decay is well-known (See Friz and Oberhauser [5]). Next, take $\alpha' \in (\alpha, 1/2)$. Then, due to the continuous embedding $G\Omega_{\alpha'}^H(\mathbf{R}^d) \hookrightarrow G\Omega_{\alpha,4m}^B(\mathbf{R}^d)$, we can easily see (8.4) holds on the Besov geometric rough path space, too.

Take $\chi \in C_b^\infty(\mathbf{R} \to \mathbf{R})$ such that $\chi = 0$ on $(-\infty, 0]$, $\chi = 1$ on $[1,\infty)$, and χ is non-decreasing. Set $G(w) = \|W^1\|_{\alpha,4m-B}^{4m}$. Then, $G \in \mathbf{D}_\infty$ and $\chi(G - R^{4m} + 1) \in \mathcal{F}_{q,r}(\{\|W^1\|_{\alpha,4m-B} \geq R\})$ for any q,r in the sense of (4.7), since $w \mapsto W$ is ∞-quasi-continuous by Corollary 5.4. By (4.8), we have

$$\mathrm{c}_{q,r}\bigl(\{\|W^1\|_{\alpha,4m-B} \geq R\}\bigr) \leq \|\chi(G - R^{4m} + 1)\|_{q,r}.$$

So, we have only to estimate the Sobolev norm on the right hand side.

First we calculate the L^q-norm.

$$\|\chi(G - R^{4m} + 1)\|_q \leq \mathbb{P}\big(\{\|W^1\|_{\alpha,4m-B} \geq (R^{4m} - 1)^{1/4m}\}\big)^{1/q} \leq e^{-(c_1/2q)R^2}$$

for sufficiently large R. It is easy to see that $D[\chi(G - R^{4m} + 1)] = \chi'(G - R^{4m} + 1)DG$. Hence,

$$\begin{aligned}\|D[\chi(G - R^{4m} + 1)]\|_q &\leq \|\chi'\|_\infty \big\| I_{\{G-R^{4m}+1\geq 0\}} \|DG\|_{\mathcal{H}} \big\|_q \\ &\leq \|\chi'\|_\infty \|DG\|_{2q} \mathbb{P}\big(\{\|W^1\|_{\alpha,4m-B} \geq (R^{4m} - 1)^{1/4m}\}\big)^{1/2q} \\ &\leq \|\chi'\|_\infty \|DG\|_{2q} \cdot e^{-(c_1/4q)R^2}\end{aligned}$$

for sufficiently large R. From this and Meyer's equivalence, there exists $c_2 = c_2(\alpha, m, q) > 0$ such that

$$\|\chi(G - R^{4m} + 1)\|_{1,q} \leq e^{-c_2 R^2}$$

for sufficiently large R. In the same way, we can estimate $D^k[\chi(G - R^{4m} + 1)]$ for any $k \in \mathbf{R}_+$ and prove the lemma for $i = 1$.

For $i = 2$, just consider

$$\tilde{G}(w) = \|W^2\|^{2m}_{2\alpha,2m-B} \quad \text{and} \quad \chi(\tilde{G} - R^{4m} + 1) \in \mathcal{F}_{q,r}(\{\|W^2\|^{1/2}_{2\alpha,2m-B} \geq R\}.$$

Then, the same argument works for this case, too. ▮

9 Proof of Theorem 7.1 (i): Uppeer estimate

The aim of this section is to prove the upper estimate in Theorem 7.1 (i). In this section, we will often use the following fact; For $f, g : (0, 1] \to [0, \infty)$, it holds that $\limsup_{\varepsilon\searrow 0} \varepsilon^2 \log(f_\varepsilon + g_\varepsilon) \leq [\limsup_{\varepsilon\searrow 0} \varepsilon^2 \log f_\varepsilon] \vee [\limsup_{\varepsilon\searrow 0} \varepsilon^2 \log g_\varepsilon]$.

[Step 1] We divide the proof into three steps. The first step is to show that

$$\lim_{R\searrow 0} \limsup_{\varepsilon\searrow 0} \varepsilon^2 \log \mu^\varepsilon_{a,a'}(B_R(X)) \leq -I(X), \qquad X \in G\Omega^B_{\alpha,4m}(\mathbf{R}^d), \tag{9.1}$$

where

$$B_R(X) = \{Y \in G\Omega^B_{\alpha,4m}(\mathbf{R}^d) \mid \|Y^i - X^i\|_{i\alpha,4m/i-B} < R^i \quad (i = 1, 2)\}.$$

First, we consider the case $a + \Phi_0(\tau_\lambda(X))^1_{0,1} \neq a'$. For simplicity we write $\tilde{a} = a + \Phi_0(\tau_\lambda(X))^1_{0,1}$. By continuity of Lyons-Itô map, there exist $\varepsilon_0 > 0$ and $R > 0$ such that $|a + \Phi_\varepsilon(\tau_\lambda(Z))^1_{0,1} - \tilde{a}| \leq |a' - \tilde{a}|/3$ holds for all $0 \leq \varepsilon \leq \varepsilon_0$ and $Z \in B_R(X)$. If we have $\mu^\varepsilon_{a,a'}(B_R(X)) = 0$ for such ε and R, then (9.1) immediately follows for this case.

Let us verify the fact $\mu^\varepsilon_{a,a'}(B_R(X)) = 0$ as follows. Let $\psi : \mathbf{R} \to [0, 1]$ be a smooth even function such that $\psi = 1$ on $[0, 1]$ and $\psi = 0$ on $[2, \infty)$ and non-increasing on $[1, 2]$.

Set $\psi(|(y_1^\varepsilon - a')/\eta|^2)$, where $\eta := |a' - \tilde{a}|/3$. Then, $\delta_{a'}(y_1^\varepsilon) = \psi(|(y_1^\varepsilon - a')/\eta|^2) \cdot \delta_{a'}(y_1^\varepsilon)$ in $\mathbf{D}_{-\infty}$. By Sugita's theorem, $\theta^\varepsilon_{a,a'}(dw) = \psi(|(a + \Phi_\varepsilon(\tau_\lambda(\varepsilon W))^1_{0,1} - a')/\eta|^2) \cdot \theta^\varepsilon_{a,a'}(dw)$, since $a + \Phi_\varepsilon(\tau_\lambda(\varepsilon W))^1_{0,1}$ is the ∞-quasi-redefinition of $y_1^\varepsilon = y^\varepsilon(1, a, w)$. Hence,

$$\begin{aligned}\mu^\varepsilon_{a,a'}(B_R(X)) &= \theta^\varepsilon_{a,a'}(\{w \in \mathcal{W} \mid \varepsilon W \in B_R(X)\}) \\ &= \int_{\mathcal{W}} I_{\{\varepsilon W \in B_R(X)\}} \cdot \psi(|(a + \Phi_\varepsilon(\tau_\lambda(\varepsilon W))^1_{0,1} - a')/\eta|^2)\theta^\varepsilon_{a,a'}(dw) = 0.\end{aligned}$$

Next we consider the case $a + \Phi_0(\tau_\lambda(X))^1_{0,1} = a'$. Note that $\|\varepsilon W^1 - X^1\|^{4m}_{\alpha,4m-B}$ and $\|\varepsilon^2 W^2 - X^2\|^{2m}_{2\alpha,2m-B}$ are in $\hat{\mathcal{C}}_{4m}$, even if $X \notin S_2(\mathcal{H})$. Since their L^2-norms are bounded in ε, so are their (q,r)-Sobolev norms for any (q,r). Note also that $\|D^r y_1^\varepsilon\|_q$ is bounded in ε for any (q,r). Set $G(u_1, \ldots, u_n) = \prod_{j=1}^n (u_j - a'_j)^+$. This is a continuous function from $\mathbf{R}^n$ to $\mathbf{R}$ with polynomial growth and satisfies $\partial_1^2 \cdots \partial_n^2 G(u) = \delta_{a'}(u)$ in the sense of Schwartz distributions.

Then, we have

$$\begin{aligned}\mu^\varepsilon_{a,a'}(B_R(X)) &= \theta^\varepsilon_{a,a'}(\{w \in \mathcal{W} \mid \varepsilon W \in B_R(X)\}) \\ &\le \mathbb{E}\Big[\prod_{i=1}^2 \psi\big(\|\varepsilon^i W^i - X^i\|^{4m/i}_{i\alpha,4m/i-B}/R^{4m}\big) \cdot (\partial_1^2 \cdots \partial_n^2 G)(y_1^\varepsilon)\Big]. \end{aligned} \tag{9.2}$$

Now we use integration by parts as in (4.2)–(4.3). Then, the right hand side of (9.2) is equal to a finite sum of the following form;

$$\sum_{j,k} \mathbb{E}\Big[F^\varepsilon_{j,k} \cdot \psi^{(j)}\big(\|\varepsilon W^1 - X^1\|^{4m}_{\alpha,4m-B}/R^{4m}\big)\psi^{(k)}\big(\|\varepsilon^2 W^2 - X^2\|^{2m}_{2\alpha,2m-B}/R^{4m}\big)G(y_1^\varepsilon)\Big]. \tag{9.3}$$

Here, $F^\varepsilon_{j,k}(w) = F_{j,k}(\varepsilon, w)$ is a polynomial in components of (i) y_1^ε and its derivatives, (ii) $\|\varepsilon^i W^i - X^i\|^{4m/i}_{i\alpha,4m/i-B}$ and its derivatives, (iii) $\tau(\varepsilon)$, which is a Malliavin covariance matrix of y_1^ε, and (iv) $\gamma(\varepsilon) = \tau(\varepsilon)^{-1}$. Note that derivatives of $\gamma(\varepsilon)$ do not appear.

By the uniform non-degeneracy of the Malliavin covariance matrix of $(y_1^\varepsilon - a)/\varepsilon$, there exists $l > 0$ such that $\gamma(\varepsilon)$ is $O(\varepsilon^{-l})$ in any L^q. In turn, this implies that there exists $l > 0$ such that $F^\varepsilon_{j,k}$ is $O(\varepsilon^{-l})$ in any L^q. (Here and in what follows, l may change from line to line.)

Take $q, q' \in (1, \infty)$ so that $1/q + 1/q' = 1$. By Hölder's inequality, the right hand side of (9.3) is dominated by

$$C\varepsilon^{-l}\mu^\varepsilon_{a,a'}\Big(\|\varepsilon^i W^i - X^i\|^{1/i}_{i\alpha,4m/i-B} \le 2^{\frac{1}{m}} R \ (i = 1, 2)\Big)^{\frac{1}{q'}} = C\varepsilon^{-l}\mu^\varepsilon_{a,a'}(\varepsilon W \in B_{2^{\frac{1}{m}}R}(X))^{\frac{1}{q'}}.$$

Here, $C = C_{q,q'} > 0$ is a constant independent of ε. By the large deviation principle of Schilder-type for the law of εW on $G\Omega^B_{\alpha,4m}(\mathbf{R}^d)$,

$$\limsup_{\varepsilon \searrow 0} \varepsilon^2 \log \mu^\varepsilon_{a,a'}(B_R(X)) \le -\frac{1}{q'} \inf\{\|h\|^2_{\mathcal{H}}/2 \mid h \in \mathcal{H}, S_2(h) \in B_{2^{1/m}R}(X)\}.$$

Letting $q' \searrow 1$, we have

$$\begin{aligned}\limsup_{\varepsilon\searrow 0} \varepsilon^2 \log \mu^{\varepsilon}_{a,a'}(B_R(X)) &\leq -\inf\{\|h\|^2_{\mathcal{H}}/2 \mid h\in\mathcal{H}, S_2(h)\in B_{2^{1/m}R}(X)\} \\ &= -\inf\{I_{Sch}(Y) \mid Y\in B_{2^{1/m}R}(X)\}.\end{aligned}$$

Since the rate function $I_{Sch} : G\Omega^B_{\alpha,4m}(\mathbf{R}^d) \to [0,\infty]$ is lower semicontinuous, the limit of the right hand side as $R \searrow 0$ is dominated by $-I_{Sch}(X)$. This proves (9.1). (Here, $I_{Sch}(S_2(h)) = \|h\|^2_{\mathcal{H}}/2$ and $I_{Sch}(X) = 0$ if X is not lying above any $h \in \mathcal{H}$.)

[Step 2] The second step is to prove the upper bound in Theorem 7.1 (i) when A is a compact set. Let $N \in \mathbf{N}_+$. For any $Y \in A$, take $R = R_{N,Y} > 0$ small enough so that

$$\limsup_{\varepsilon\searrow 0} \varepsilon^2 \log \mu^{\varepsilon}_{a,a'}(B_R(Y)) \leq \begin{cases} -N & (\text{if } I(Y)=\infty), \\ -I(Y)+N^{-1} & (\text{if } I(Y)<\infty).\end{cases}$$

The union of such open balls over $Y \in A$ covers the compact set A. Hence, there are finitely many $Y_1, \ldots, Y_k$ such that $A \subset \cup_{i=1}^k B_{R_i}(Y_i)$, where $R_i = R(N, Y_i)$. By using the remark in the beginning of this section, we see that

$$\begin{aligned}\limsup_{\varepsilon\searrow 0} \varepsilon^2 \log \mu^{\varepsilon}_{a,a'}(A) &\leq (-N) \vee \max\{-I(Y_i) + N^{-1} \mid 1\leq i\leq k,\ I(Y_i)<\infty\} \\ &\leq (-N) \vee \big(-\inf_{h\in A\cap\mathcal{K}^{a,a'}} \|h\|^2_{\mathcal{H}}/2 + N^{-1}\big).\end{aligned}$$

Letting $N \to \infty$, we obtain

$$\limsup_{\varepsilon\searrow 0} \varepsilon^2 \log \mu^{\varepsilon}_{a,a'}(A) \leq -\inf\{\|h\|^2_{\mathcal{H}}/2 \mid h\in\mathcal{K}^{a,a'},\ S_2(h)\in A\}.$$

Thus, we have obtained the upper estimate for the compact case.

[Step 3] In the final step, we verify the case when A is a closed set. Take $\alpha' \in (\alpha, 1/2)$ so that (α', m) still satisfies (5.19). As in the proof of Proposition 8.4, let $B'_R(H) = T_h(B_R)$ be the "ball" of radius $R > 0$ centered at H in $(\alpha', 4m)$-Besov geometric rough path space. (We will use the same symbol by abusing notations.) Note that this is precompact in $G\Omega^B_{\alpha,4m}(\mathbf{R}^d)$ by Proposition 3.2.

Choose any $\bar{h} \in \mathcal{K}^{a,a'} \cap \mathcal{W}^*$. For sufficiently large $R > 0$,

$$\begin{aligned}\mu^{\varepsilon}_{a,a'}(B'_R(\bar{H})^c) &= \int_{\mathcal{W}} I_{T_{\bar{h}}(B_R)^c}(\varepsilon W)\theta^{\varepsilon}_{a,a'}(dw) \\ &= \varepsilon^{-n}\int_{\mathcal{W}} \exp\big(-\frac{\langle \bar{h}, w\rangle}{\varepsilon} - \frac{\|\bar{h}\|^2_{\mathcal{H}}}{2\varepsilon^2}\big) I_{\{\|W^i\|^{1/i}_{i\alpha',4m/i-B}<R/\varepsilon\ (i=1,2)\}^c}\nu^{\varepsilon}(dw) \\ &\leq \varepsilon^{-n} e^{\|\bar{h}\|^2_{\mathcal{H}}/2\varepsilon^2}\mathbb{E}\Big[\exp\big(-\frac{\langle 2\bar{h}, w\rangle}{\varepsilon} - \frac{\|2\bar{h}\|^2_{\mathcal{H}}}{2\varepsilon^2}\big)\delta_0\big(\frac{y^{\varepsilon,\bar{h}}_1 - a'}{\varepsilon}\big)\Big]^{1/2} \\ &\quad \times \Big\{\sum_{i=1,2} \nu^{\varepsilon}(\{\|W^i\|^{1/i}_{i\alpha',4m/i-B} \geq R/\varepsilon)\Big\}^{1/2}.\end{aligned}$$

As we proved in (the proof of) Proposition 8.4 and Lemma 8.5, the last factor above is dominated by $2e^{-c_1R^2/\varepsilon^2}$ for some constant $c_1 > 0$. Since $\|\delta_0((y_1^{\varepsilon,\bar{h}} - a')/\varepsilon)\|_{q',-r}$ is bounded in ε for some q', r, it suffices to estimate the (q,r)-Sobolev norm of the Cameron-Martin density function $\exp(-\langle 2\bar{h}, w\rangle/\varepsilon - \|2\bar{h}\|^2_{\mathcal{H}}/2\varepsilon^2)$. Using Hölder's inequality, we can easily see that its L^q-norm is dominated by $e^{2(q-1)\|\bar{h}\|^2_{\mathcal{H}}/\varepsilon^2}$. We can also estimate L^q-norms of its derivatives and, by using Meyer's equivalence, we have

$$\Big\|\exp\big(\frac{\langle 2\bar{h}, w\rangle}{\varepsilon} - \frac{\|2\bar{h}\|^2_{\mathcal{H}}}{2\varepsilon^2}\big)\Big\|_{q,r} \leq c_2\varepsilon^{-c_3}(1 + \|h\|_{\mathcal{H}})^{c_4} e^{2(q-1)\|\bar{h}\|^2_{\mathcal{H}}/\varepsilon^2}$$

for some positive constants $c_j = c_{j,q,r}$ $(j = 2, 3, 4)$, which are independent of ε. Therefore, we have

$$\limsup_{\varepsilon\searrow 0} \varepsilon^2 \log \mu^\varepsilon_{a,a'}(B'_R(\bar{H})^c) \leq -c_1R^2 + (q - \frac{1}{2})\|\bar{h}\|^2_{\mathcal{H}}.$$

Note that the right hand side can be made arbitrarily small if we take R large enough.

Now let A be a closed set such that $A \cap S_2(\mathcal{K}^{a,a'}) \neq \emptyset$. Take $R > 0$ so that

$$-c_1R^2 + (q - \frac{1}{2})\|\bar{h}\|^2_{\mathcal{H}} < -\inf_{h\in A\cap\mathcal{K}^{a,a'}} \|h\|^2_{\mathcal{H}}/2.$$

Decompose $A = \overline{B'_R(\bar{H})} \cap A + (\overline{B'_R(\bar{H})})^c \cap A$ into a union of two disjoint sets. The second set is included in $B'_R(\bar{H})^c$. The first set in the disjoint union is compact and, due to the previous step, we have

$$\begin{aligned}\limsup_{\varepsilon\searrow 0} \varepsilon^2 \log \mu^\varepsilon_{a,a'}(\overline{B'_R(\bar{H})} \cap A) &\leq -\inf\{\|h\|^2_{\mathcal{H}}/2 \mid h \in \mathcal{K}^{a,a'},\, S_2(h) \in A \cap \overline{B'_R(\bar{H})}\} \\ &\leq -\inf\{\|h\|^2_{\mathcal{H}}/2 \mid h \in \mathcal{K}^{a,a'},\, S_2(h) \in A\}.\end{aligned}$$

By the remark in the beginning of this section, we obtain the upper estimate in this case. The case $A \cap S_2(\mathcal{K}^{a,a'}) = \emptyset$ can be done in a similar way and is actually easier. This completes the proof of Theorem 7.1.

References

[1] Aida, S.; Vanishing of one-dimensional L^2-cohomologies of loop groups. J. Funct. Anal. 261 (2011), no. 8, 2164–2213.

[2] Dembo, A., Zeitouni, O.; Large deviations techniques and applications. Second edition. Springer-Verlag, New York, 1998.

[3] Dereich, S.; Rough paths analysis of general Banach space-valued Wiener processes. J. Funct. Anal. 258 (2010), no. 9, 2910–2936.

[4] Friedman, A.; Stochastic differential equations and applications. Vol. 2. Academic Press, New York-London, 1976.

[5] Friz, P.; Oberhauser, H., A generalized Fernique theorem and applications. Proc. Amer. Math. Soc. 138 (2010), no. 10, 3679–3688.

[6] Friz, P.; Victoir, N.; Approximations of the Brownian rough path with applications to stochastic analysis. Ann. Inst. H. Poincaré Probab. Statist. 41 (2005), no. 4, 703–724.

[7] Friz, P.; Victoir, N.; Large deviation principle for enhanced Gaussian processes. Ann. Inst. H. Poincaré Probab. Statist. 43 (2007), no. 6, 775–785.

[8] Friz, P.; Victoir, N.; Multidimensional stochastic processes as rough paths. Cambridge University Press, Cambridge, 2010.

[9] Higuchi, Y.; Master Theses (in Japanese), Graduate School of Engineering Sciences, Osaka University, 2006.

[10] Hsu, P.; Brownian bridges on Riemannian manifolds. Probab. Theory Related Fields 84 (1990), no. 1, 103–118.

[11] Ikeda, N., Watanabe, S.; Stochastic differential equations and diffusion processes. Second edition. North-Holland Publishing Co., Amsterdam; Kodansha, Ltd., Tokyo, 1989.

[12] Inahama, Y.; Quasi-sure existence of Brownian rough paths and a construction of Brownian pants. Infin. Dimens. Anal. Quantum Probab. Relat. Top. 9 (2006), no. 4, 513–528.

[13] Inahama, Y., Kawabi, H.; Large deviations for heat kernel measures on loop spaces via rough paths. J. London Math. Soc. (2) 73 (2006), no. 3, 797–816.

[14] Ledoux, M.; Qian, Z.; Zhang, T.; Large deviations and support theorem for diffusion processes via rough paths. Stochastic Process. Appl. 102 (2002), no. 2, 265–283.

[15] Lyons, T.; Differential equations driven by rough signals. Rev. Mat. Iberoamericana 14 (1998), no. 2, 215–310.

[16] Lyons, T., Qian, Z.; System control and rough paths. Oxford University Press, Oxford, 2002.

[17] Malliavin, P.; Stochastic analysis. Springer-Verlag, Berlin, 1997.

[18] Malliavin, P.; Nualart, D.; Quasi-sure analysis of stochastic flows and Banach space valued smooth functionals on the Wiener space. J. Funct. Anal. 112 (1993), no. 2, 287–317.

[19] Millet, A., Sanz-Solé, M.; Large deviations for rough paths of the fractional Brownian motion. Ann. Inst. H. Poincaré Probab. Statist. 42 (2006), no. 2, 245–271.

[20] Ren, J.; Analyse quasi-sûre des équations différentielles stochastiques. Bull. Sci. Math. 114 (1990), no. 2, 187–213.

[21] Shigekawa, I.; Stochastic analysis. Translations of Mathematical Monographs, 224. Iwanami Series in Modern Mathematics. American Mathematical Society, Providence, RI, 2004.

[22] Sugita, H.; Positive generalized Wiener functions and potential theory over abstract Wiener spaces. Osaka J. Math. 25 (1988), no. 3, 665–696.

[23] Takanobu, S.; Watanabe, S.; Asymptotic expansion formulas of the Schilder type for a class of conditional Wiener functional integrations. Asymptotic problems in probability theory: Wiener functionals and asymptotics (Sanda/Kyoto, 1990), 194–241, Pitman Res. Notes Math. Ser., 284, Longman Sci. Tech., Harlow, 1993.

[24] Watanabe, S.; Analysis of Wiener functionals (Malliavin calculus) and its applications to heat kernels. Ann. Probab. 15 (1987), no. 1, 1–39.

[25] Watanabe, S.; Itô calculus and Malliavin calculus. Stochastic analysis and applications, 623–639, Abel Symp., 2, Springer, Berlin, 2007.